PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :		(11) International Publication Number:
C12Q 1/68	A1	WO 92/16657
		(43) International Publication Date: 1 October 1992 (01.10.92)

(21) International Application Number: PCT/US92/01691

(22) International Filing Date: 12 March 1992 (12.03.92)

(30) Priority data:
669,568 13 March 1991 (13.03.91) US

(60) Parent Application or Grant
(63) Related by Continuation
US 669,568 (CIP)
Filed on 13 March 1991 (13.03.91)

(71) Applicant *(for all designated States except US):* E.I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only) :* LIVAK, Kenneth, J. [US/US]; 2418 Shellpot Drive, Wilmington, DE 19803 (US). RAFALSKI, Jan, Antoni [US/US]; 2028 Longcome Drive, Wilmington, DE 19810 (US). SHEPHERD, Nancy, Faye, Stacy [US/US]; 210 Partridge Way, Kennett Square, PA 19348 (US).

(74) Agent: GALLEGOS, R., Thomas; E.I. du Pont de Nemours and Company, Legal/Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).

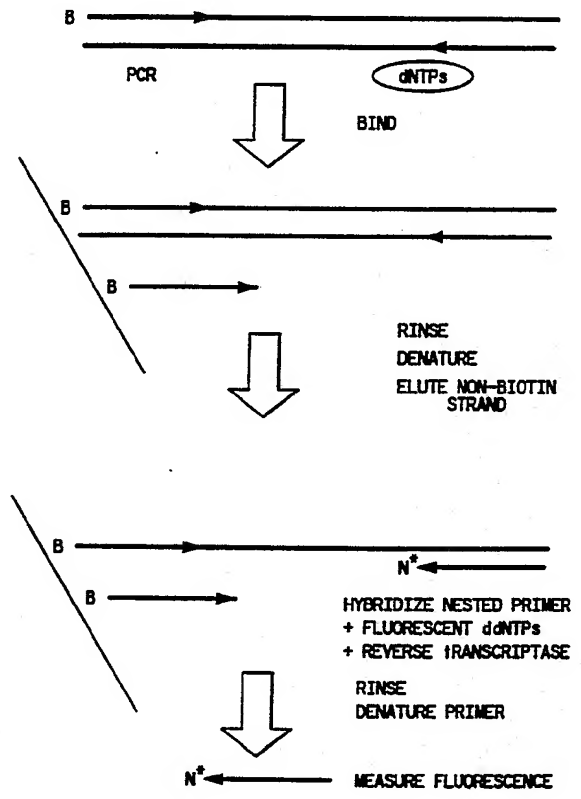
(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), MC (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.

(54) Title: METHOD OF IDENTIFYING A NUCLEOTIDE PRESENT AT A DEFINED POSITION IN A NUCLEIC ACID

(57) Abstract

A method is described for identifying a nucleotide at a defined point on a nucleic acid sequence. An oligonucleotide probe is annealed to a target nucleotide sequence of the nucleic acid sample at a point immediately adjacent and 3′ to the nucleotide of interest. The probe is then extended in the direction of the nucleotide of interest in a reaction medium containing at least one chain terminating nucleotide triphosphate (ATP, GTP, TTP and CTP). The nucleotide of interest is complementary to the labeled nucleotide incorporated into the primer by the extension reaction.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	Finland	ML	Mali
AU	Australia	FR	France	MN	Mongolia
BB	Barbados	GA	Gabon	MR	Mauritania
BE	Belgium	GB	United Kingdom	MW	Malawi
BF	Burkina Faso	GN	Guinea	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IE	Ireland	RO	Romania
CA	Canada	IT	Italy	RU	Russian Federation
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		
ES	Spain	MG	Madagascar		

TITLE

METHOD OF IDENTIFYING A NUCLEOTIDE PRESENT AT A DEFINED POSITION IN A NUCLEIC ACID

FIELD OF THE INVENTION

This invention relates to a rapid, convenient process to identify a nucleotide present at a specific position in a nucleic acid chain (DNA or RNA) of a biological sample.

BACKGROUND OF THE INVENTION

The nucleic acid content of any organism is the essence of that organism, and differences in the nucleic acid are known to be of primary importance in distinguishing one from another. The science of genetics is based on the identification and characterization of differences in nucleic acid sequence. These differences, or polymorphisms, are often termed "mutations" and may be due to nucleotide substitution, insertion or deletion. Thus, many techniques have been developed to compare homologous segments of DNA or RNA to determine if the segments are identical or if they differ at one or more nucleotides. Identification of genetic polymorphisms is useful for genetic diagnoses in medicine, identification of individuals in forensic science, identification of pathogenic organisms, construction of genetic polymorphism maps for locating genes important in disease and in agriculture and for breeding of plants and animals.

The most definitive method for comparing DNA segments is to determine the complete nucleotide sequence of each segment. Examples of how sequencing has been used to study mutations in human genes are included in the publications of Engelke, et. al., Proc.

Natl. Acad. Sci. U.S.A. 85:544-548 (1988) and Wong, et al., Nature 330:384-386 (1987). At the present time, it is not practical to use extensive sequencing to compare more than just a few DNA segments, because the effort required to determine, interpret, and compare sequence information is time-consuming and costly. Development of automated sequencing technology; Innis et al., WO 9003442; Prober et al., EPO 252638; Rosenthal et al., J. DNA Sequencing and Mapping 1:63-71 (1990); Bauer, Nucleic Acids Res. 18:879-884 (1990); has made the process more efficient and has substituted fluorescent reporters in place of radioactivity; however, these methods involve a time-consuming separation of reaction products using a polyacrylamide gel, and the nucleotide sequence information at any one point in a nucleotide chain is one of several hundred such pieces of information. That is, if one is only interested in determining the nucleotide at a specific point in the polynucleotide chain (e.g. a position known from previous genetic analysis to be involved in a disease phenotype), that information must be retrieved from a larger volume of data.

For genetic mapping purposes, the most commonly used screen for DNA polymorphisms consists of digesting DNA with restriction endonucleases and analyzing the resulting fragments by means of Southern blots, a method known as "Restriction Fragment Length Polymorphism" or RFLP mapping, as described by Botstein, et al., Am. J. Hum. Genet. 32:314-331 (1980); White, et al., Sci. Am. 258:40-48 (1988). Mutations that affect the recognition sequence of the endonuclease will preclude enzymatic cleavage at that site, thereby altering the cleavage pattern of that DNA. DNAs are compared by looking for differences in restriction fragment lengths. RFLP

detection in a genomic DNA sample is very labor intensive for it requires preliminary steps of genomic DNA isolation, restriction, gel electrophoresis and Southern transfer steps, before hybridization to a probe that is generally radioactively labeled for sensitive detection of homologous sequences. A major problem associated with RFLP detection is the necessity of the polymorphism to affect cleavage with a restriction endonuclease, therefore many mutations cannot be detected with this method (Jeffreys, Cell 18:1-18, 1979). More importantly, although RFLP and several other methods in the prior art (e.g. Wallace et al., Nucl. Acids Res. 9:879-894, 1981 or Saiki, et al., U.S. Pat. No. 4,683,194 or Kornher et al., U.S. Pat. No. 4,879,214) are useful for finding polymorphisms in DNA, they do not elucidate the exact nature of the nucleotide present at a specific position on the nucleic acid sequence. In some applications, such as prenatal diagnosis, knowledge of which nucleotide is present at a given position is extremely important, since some nucleotide changes do not alter the coding capacity of a gene and are therefore "silent" with respect to phenotype. Those techniques that elucidate the nature of the nucleotide present are discussed below.

Many techniques designed to elucidate the nature of a nucleic acid polymorphism involve hybridization with a polynucleotide probe, a portion of which is complementary to the nucleotide position(s) of interest. A target sequence that is perfectly complementary to the probe can be distinguished from a target that differs by as little as a single nucleotide in a variety of ways. A technique involving amplification and mismatch detection (AMD), described by Montandon et al., Nucl. Acids Res. 17:3347-3358, 1989, utilizes amplification of

the DNA region of interest from two samples, followed by denaturation and reannealing to form homo- and heteroduplexes between DNA molecules of the two samples. Any nucleotide position that is different between the two amplified DNAs (i.e. nucleotide mismatch in the heteroduplex) can be identified with the use of hydroxylamine and osmium tetroxide to modify mispaired cytosines and thymines, respectively, followed by piperidine-catalysed cleavage of the modified heteroduplexes, and subsequent gel electrophoresis to identify cleavage products. Although this technique and the analogous technique of EP0329311 to Campbell and Cotton, are useful for both detecting and identifying all point mutations within a nucleic acid segment, they share some of the same, serious disadvantages of the chemical degradation method for DNA sequencing by Maxam and Gilbert (Proc. Natl. Acad. Sci. 74:560, 1977). These techniques require dangerous chemicals that modify and cleave nucleic acids, they involve several different chemical reactions, and require a time-consuming gel assay.

Several techniques that utilize an oligonucleotide probe designed to overlap the position of interest at the 3'prime terminus are reviewed below. The technique described in Landegren, et al., Science 241:1077-1080 (1988) uses an enzymatic detection of polymorphisms. For this method, oligonucleotide probes are constructed in pairs such that their junction corresponds to the specific nucleotide site which is of interest. These oligonucleotides are then hybridized to the DNA being analyzed. Base pair mismatch between either oligonucleotide and the target DNA at the junction location prevents the efficient joining of the two oligonucleotide probes by DNA ligase. The Ligation

Amplification Reaction (LAR) as reported by Wu and Wallace, Genomics 4:560-569, (1989) and Wallace and Skolnick (WO 89/10414) is also dependent upon ligation of oligonucleotides whose 3-prime ends include the nucleotide position of interest. They demonstrate that four pairs of oligonucleotides that are complementary to the upper and lower strand of the target DNA will be exponentially amplified only if there is perfect complementarity between the oligonucleotides and the target DNA. The patent of Vary et al., U.S. Pat. No. 4,851,331, also depends upon an enzymatic reaction that requires one end of the oligonucleotide probe to form a perfect, complementary matched basepair with the target nucleotide sequence. As in the examples above, an oligonucleotide probe is designed such that the 3-prime end of the complementary probe includes the specific nucleotide position of interest. After annealing this oligonucleotide probe to the template DNA, a polymerase that replicates nucleic acid strands in a template directed fashion is used to incorporate modified nucleotides into a newly synthesized strand. If the 3-prime end of the oligonucleotide probe did not contain a nucleotide complementary to the target nucleotide sequence, then the polymerase cannot begin the replication process. The amount of incorporation is a measure of the amount of the specific template DNA in the biological sample. This same principle of utilizing a polymerase to discriminate whether there is a mismatched base at the 3-prime end of the primer was also recently combined with the PCR to give an exponential rather than linear increase of the reaction products in a process called Allele-specific Polymerase Chain Reaction (ASPCR) (Wu et al., Proc. Natl. Acad. Sci. 86:2757-2760, 1989). In their example, the

reaction products were run on an agarose gel and detected by ethidium bromide staining. However, fluorescently-labeled oligonucleotide primers may also be used for detection in ASPCR (e.g. Chehab and Kan, Proc. Natl. Acad. Sci. 86:9178-9182, 1989). The patent of Caskey and Gibbs (EPO333465A2) for a process involving Competitive Oligonucleotide Priming (COP) is essentially the same. In COP, two differentially labeled oligonucleotide primers that differ at the 3' nucleotide and overlap the position of interest are present in the same, rather than in separate reactions, and thus compete for template molecules in the hybridization reaction.

A general problem shared by the various techniques mentioned in the preceding paragraph is that the difference in duplex stability of a perfectly vs nonperfectly matched oligonucleotide to its target DNA is dependent upon the length and sequence of the oligonucleotide. Therefore, regardless of the assay method, a different set of emperically determined experimental reaction conditions may be required in order to assay different genomic loci for a polymorphism. Secondly, since their assay is a +/- assay (a reaction product should be formed or absent), it is necessary to perform several assays on a single target DNA so that an inference can be made concerning the nature of the nucleotide at a given position. This is especially important since many organisms are diploid, polyploid, or have several copies of a given sequence such that an individual may be polymorphic at any one nucleotide position (e.g. heterozygous). Several reactions utilizing oligonucleotide probes differing in sequence only at the 3-prime end may give a reaction product in such a situation. Finally, this

same result can occur if the assay method is too sensitive such that even inefficient ligation or replication is detected as a positive signal. Misincorporation of nucleotide substrate, well documented in the literature for polymerases; Ricchetti and Buc, The EMBO J. 9:1583-1593, (1990); or template-independent ligation products due to blunt end ligation; Hayashi et al., Nucl. Acids Res. 14:7617-7631, (1986); can lead to a false signal if not adequately suppressed in the reaction. Such misincorporation is especially apparent when the correct, complementary nucleotide substrate is absent from the reaction. The polymerase chain reaction is quite dependent upon products generated during the first few rounds of amplification. The difficulty of devising conditions that totally suppress amplification by the primer that contains a mismatched base at the 3' end is also documented in Chehab and Kan; Proc. Natl. Acad. Sci. 86:9181 (1989); where fluorescence values as high as 0.8 were considered negative for they were less than 1.0, while all values above 1.0 were considered positive for amplification (even values as low as 1.4 were considered positive).

In the technique described in Mundy, U.S. Pat. No. 4,656,127, specific mutations can be detected by first hybridizing a labeled DNA probe to the target nucleic acid in order to form a hybrid in which the 3' end of the probe is positioned adjacent to the specific base being analyzed. Then, a DNA polymerase is used to add a nucleotide analog, such as a thionucleotide, to the probe strand, but only if the analog is complementary to the specific base being analyzed. Finally, the probe-target hybrid is treated with exonuclease III. If the nucleotide analog has been incorporated, the labeled probe is protected from nuclease digestion. Absence of

a labeled probe indicates that the analog and the specific base being analyzed were not complementary. For this technique, the nucleotide analog may be given as the sole substrate for incorporation, or it may be added as one of a maximum of three nucleotide substrates. It is critical to note that all four bases cannot be given as substrate, since this would allow chain elongation to occur until eventually the modified nucleotide analog would be complementary and thus incorporated. This limits the utility of the assay to certain DNA sequences that can have the appropriate positioning of primers. For it is well known that if the correct base is not supplied in a reaction that a "wobble" base pair between G and T will occur to a significant degree (Boosalis et al., J. of Biol. Chem. 262:14689-14696, 1987). Unlike the reaction described in Mundy, the reaction of the present invention can include all four nucleotide substrates, thus leading to a lower incorporation of non-complementary nucleotide.

A similar technique described by Sokolov; Nucl. Acids. Res. 18:3671, (1989); utilizes an unlabeled oligonucleotide probe that is perfectly complementary to the target sequence and again positioned with the 3-prime end of the probe adjacent to the nucleotide position of interest. Four separate reactions are performed, each with only one of the four radioactively labeled dNTP's (dATP, dGTP, dCTP, dTTP) supplied as substrate for a primer extension reaction using Taq-polymerase. Reaction products are run on a standard polyacrylamide sequencing gel, and the level of nucleotide incorporation in each of the four reactions is monitored by autoradiography. In this manner, the specific nucleotide at the position of interest is indicated.

The present invention solves several problems inherent in the Sokolov method. (1) This invention is not dependent upon radioactive substrates nor the time-consuming monitoring of the assay via a polyacrylamide gel and subsequent autoradiography. (2) This invention uses chain-terminating nucleotides as substrates in the reaction, therefore preventing incorporation of several of the same nucleotide in the primer extension product if there are several of the same nucleotides present in a row on the template. (3) The analysis of Sokolov required four separate reactions whereas the present invention would need only one reaction to gain the same amount of information. (4) As mentioned above, if the correct, complementary nucleotide substrate is not present in the reaction, then significant misincorporation can occur in the Sokolov reaction. Misincorporation is substantially prevented in the present invention.

Automation of the ASPCR reaction was described but not demonstrated in Wu et al., Proc. Natl. Acad. Sci. 86:2757-2760, (1989), and again by Chehab and Kan, Proc. Natl. Acad. Sci. 86:9178-9182, (1989), for fluorescent ASPCR. In both cases, each ASPCR reaction is performed using one biotin-labeled primer and one fluorescently-labeled primer. The biotinylated, double-stranded amplification products are then separated from unincorporated fluorescent primer using streptavidin coated magnetic beads. The color of the amplified DNA would then be determined fluorometricly through a fiber optic bundle, or alternatively, by separation and detection on a sequencing gel as is currently performed for DNA sequencing using fluorescently labeled primers. The differences between this method and that of the present invention are significant. Most importantly,

the reaction products due to amplification from the "wrong" primer are also biotinylated and will be retained in the capture process. Furthermore, these products due to misincorporation will electrophorese to the same position in the sequencing gel, thus interfering with the analysis. Finally, the size of the amplified products will in general be larger than an oligonucleotide, thus requiring a longer time for gel electrophoresis and detection than that of the present invention.

In order to be useful for a wide variety of applications, a technique that identifies a nucleotide at a given position in a nucleic acid sample should be fast, simple, reliable, and avoid radioactive or nucleic acid cleaving compounds. The currently available detection techniques discussed above are deficient in one or more of these areas. All of these problems are overcome by the present invention.

The process of the present invention exploits some of the same principles and advantages as described for sequencing of nucleic acids using fluorescent dideoxynucleotide substrates (Prober et al., EPO 252638; Mitchell and Merril WO 89/12063; Innis et al., WO 9003442). It is, however, different from that process in that one of the components of the process namely the dNTP substrates allowing multiple lengths of primer elongation product is missing in this invention. The method of the present invention generally requires a knowledge of the nucleic acid sequence in the region of interest, but it does not require the mutation to be at a restriction enzyme cleavage site. The method is capable of giving unambiguous results. In the preferred forms, it does not require tedious preliminary steps that characterize prior methods and indeed lends itself

to automation due to the lack of centrifugation steps and the ability to quickly assay reaction products without a gel separation step. If a gel separation assay is used, then multiplexing of samples based on differences in length of the probe oligonucleotide is possible. Alternatively, the probe oligonucleotide of several reactions can be of the same length, but loaded at different times after pausing the electrophoresis run.

SUMMARY OF THE INVENTION

The present invention provides a process for identifying the nucleotide present at a specific position in a nucleic acid sequence. It is based upon the selective attachment of one of four chain-terminating nucleotides, that are detectably labeled and distinguishable, onto a probe in a complementary, template dependent fashion. The probe is designed to selectively hybridize to a target nucleotide sequence and oriented such that a one nucleotide extension of the probe, usually in the 3-prime direction, will base pair to the nucleotide position of interest. The oligonucleotide probe, the nucleic acid containing the target nucleotide sequence, or both, may contain a site for specific immobilization to facilitate separation from unincorporated nucleotides and primers, such that the labeled nucleotides incorporated into the reaction product can be measured without use of a gel system such as agarose or acrylamide.

Thus the present invention provides a method for the identification of the nucleotide present at a single, defined position in the nucleic acid which comprises the following steps:

(a) contacting a nucleic acid analyte with a probe oligonucleotide of sufficient length and appropriate

sequence under conditions sufficient for the probe to bind preferentially to a target nucleotide sequence and form a hybrid having a double-stranded portion including the 3-prime end of the probe. The nucleotide position of interest is the first base of the nucleic acid analyte which extends in a 3' to 5' direction beyond the 3' end of the probe nucleotide sequence and is immediately adjacent to the hybrid formed (Figure 1).

(b) extending the probe oligonucleotide strand of the hybrid beyond its 3' end in the 5' to 3' direction by enzymatic addition of a detectably labeled, chain-terminating nucleotide which is complementary to the nucleotide position of interest;

(c) detecting whether a detectably-labeled chain-terminating nucleotide has been incorporated to determine which of the complementary nucleotide base(s) is present at the target nucleotide position of interest.

(d) identifying the nucleotide of interest as the nucleotide complementary to the incorporated chain-terminating nucleotide.

The present invention also provides a kit. The kit includes reagents in packaged form. The package may include extension primers, a probe polynucleotide, chain-terminating nucleotides and extension enzymes. The package may include attachment moieties and solid supports. Any of the reagents may include attachment moieties. A package may include any combination of reagents as necessary for a particular purpose. A package may include an insert such as a standard or direction for handling specific reagents.

DESCRIPTION OF FIGURES

Figure 1, comprising Figures 1a-1h, illustrate in various schematic forms, the location of various components of the process of this invention.

Figure 1a illustrates an analyte strand (An) which contains the nucleotide position of interest (N), the identity of which is to be determined by the assay. A target nucleotide sequence (TNS) immediately 3' of, but not including the nucleotide position of interest is illustrated. A double strand nucleic acid region forms when a probe binds to analyte strand An by complementary base pairing to the target nucleotide sequence TNS.

Figure 1b illustrates the incorporation of a chain terminating nucleotide (N*) complementary to the nucleotide of interest (N) after contacting the double stranded region in Figure 1a with a polymerase capable of primer extension. (The * in this and subsequent figures is used to illustrate a detectable label attached to the nucleotide).

Figure 1c illustrates the same features as Figure 1a, but with a specific example showing the nucleotide of interest as a thymidine (T).

Figure 1d illustrates the same features as Figure 1b, but using the same specific example as Figure 1c, to show the result of enzymatic incorporation of a detectably labeled adenosine at the 3' terminus of the probe as the nucleotide complementary to the nucleotide of interest (thymidine).

Figure 1e illustrates the incorporation of a detectably labeled guanosine at the 3' terminus of the probe and complementary to the nucleotide of interest (cytidine).

Figure 1f illustrates the use of another analyte strand for the assay (the complementary strand of the

analyte strand shown in Figure 1c). As in prior figures, the target nucleotide sequence is chosen to be immediately 3' of the nucleotide of interest (in this example shown as an adenosine), and the probe is complementary to the target nucleotide sequence.

Figure 1g illustrates the incorporation of a detectably labeled thymidine at the 3' terminus of the probe and complementary to the nucleotide of interest (adenosine).

Figure 1h illustrates the incorporation of a detectably labeled cytidine at the 3' terminus of the probe and complementary to the nucleotide of interest (guanosine).

Figure 2 illustrates an example of steps that can be used in the practice of this invention when it is desired to use a nucleic acid strand immobilized on a solid support as the analyte strand.

Figure 3 illustrates an example of steps that can be used in the practice of this invention when it is desired to use a nucleic acid strand in solution as the analyte strand.

Figure 4 illustrates identification of the nucletide of interest (N) when it is located at the 5' terminus of the analyte strand.

Figure 5 illustrates identification of a difference between two analyte strands, when that difference is part of a nucleotide insertion or deletion.

Figure 6 illustrates how the number of assays must increase if the number of distinguishably labeled chain terminating nucleotides in the reaction are decreased, if the identity of all possible nucleotides at the position of interest is to be determined.

Figure 7 illustrates the output signals obtained when the detectably labeled substrates used in the

examples are detected with the Genesis 2000 DNA analysis system.

Figure 7a illustrates the output signal [ratio of the green line to red line peak height, +/- one standard deviation] of data obtained as in Figure 7b and 7c for each of the four detectably labeled nucleotides used in the examples detected either through a gel or through a capillary.

Figure 7b is representative data showing the position and relative peak heights of the two photomultiplier tube signals (red and green lines) when SF-ddGTP-505 or SF-ddTTP-526 are electrophoresed through a urea-polyacrylamide slab gel mounted on the Genesis 2000.

Figure 7c illustrates the output signal obtained when SF-ddGTP-505 or SF-ddCTP-519 are each passed four times (therefore four peaks) through an empty capillary mounted for detection on the Genesis 2000 unit.

Figure 8 illustrates the double stranded portion of the mouse RNA polymerase II gene that was amplified using PCR primer 1 and PCR primer 2, as well as the position and sequence of the various oligonucleotide probes used in Examples 1-5.

Figure 9 illustrates the sequence of the Wildtype and the Mutant allele of the RNA polymerase II gene between nucleotides 5395 and 5454, with the difference between the two alleles indicated by boldface type at position 5430.

Figure 10 illustrates the data obtained in Example 1: Incorporation of either labeled SF-ddATP-512, SF-ddGTP-505, or both in approximately equal amounts, when probe A is used on nucleic acid samples known to be either Wildtype, Mutant, or Heterozygous at nucleotide position 5430 of the RNA polymerase II gene.

Figure 11 illustrates the data obtained in Example 2: Incorporation of either labeled SF-ddTTP-526 or SF-ddCTP-519 when probe B is used on nucleic acid samples known to be either Wildtype or Mutant at nucleotide position 5430 of the RNA polymerase II gene.

Figure 12 illustrates the data obtained in Example 3: The level of misincorporation that may occur if the correct, complementary nucleotide is not included in the reaction.

Figure 13 illustrates the data obtained in Example 4: Correct incorporation of SF-ddTTP-526 in the Wildtype allele and SF-ddCTP-519 in the Mutant allele in the presence of all four, distinguishably labeled ddNTPs in the reaction at lower nucleotide concentrations than in previous examples.

Figure 14 illustrates the double stranded nucleic acid region of the Wildtype A1 gene of Zea mays that is amplified using PCR primers A and B, with primer C indicating the sequence and position of the oligonucleotide probe used in Examples 5-6.

Figure 15 illustrates the data of Example 5: Correct incorporation of SF-ddGTP-505 when a mutant allele, a-dt, of the maize A1 gene (carrying a G-C base pair rather than C-G base pair) is assayed.

Figure 16 illustrates the data of Example 6: Use of a modified T7 polymerase (Sequenase), to give the correct incorporation of approximately equal amounts of SF-ddCTP and SF-ddGTP in a sample heterozygous for the maize allele.

DETAILED DESCRIPTION OF THE INVENTION

The present invention discloses a process for identifying the nucleotide present at defined nucleotide position in a nucleic acid sequence. This process has utility as a rapid, convenient means to genotype a

biological sample with respect to specific, nucleic acid sequence information (e.g. nucleotide positions correlated with phenotypic differences among individuals between species or in tissues). Single base pair mutations such as transitions, transversions, insertion, deletion as well as more complex rearrangements can be assayed using the method of the present invention if the appropriate oligonucleotide probe is designed (Figure 5).

The presence of a target nucleic acid in a biological sample may be detected generally as the presence or absence of an incorporated nucleotide. Individual nucleotides located at selected sites in the nucleic acid sample may also be identified. The method presented here is generally applicable to all nucleic acid sequences (DNA or RNA), whether they are single or double stranded, as long as the target nucleic acid strand is of sufficient length to form a hybrid with a complementary, oligonucleotide probe. Any source of nucleic acid, in purified or nonpurified form can be utilized as the starting nucleic acid or acids, if it contains, or is suspected of containing, the target nucleic acid sequence. The target nucleic acid can be only a fraction of a larger molecule or can be present initially as a discrete molecule. Additionally, the target nucleic acid may constitute the entire nucleic acid or may be a fraction of a complex mixture of nucleic acids.

The method of this invention requires formation of a hybrid between an oligonucleotide primer (referred to herein as the oligonucleotide probe) and the target nucleic acid sequence. Probes of relatively short length (e.g. 10-100 nucleotides) are preferred in that they can be chemically synthesized. The probe can consist of

DNA, RNA, a contiguous DNA-RNA polynucleotide, or a nucleic acid chain containing one or more modified nucleotides. Under the appropriate conditions the probe should selectively form a hybrid with the target nucleotide sequence. The probe of this invention terminates one nucleotide prior to the position of interest such that the first nucleotide to be added to the 3' terminus of the oligonucleotide probe in a template-dependent, primer extension reaction will be a nucleotide complementary to the nucleotide position of interest (Figure 1).

In practice, the specific target nucleic acid of a biological sample must be present in sufficient quantity such that hybrid molecules formed with the probe oligonucleotide are detectable by the label incorporated onto the probe.

Some samples may contain a sufficient number of target nucleic acid strands, but other samples may not. For these, molecular cloning of a region surrounding the nucleotide position of interest would suffice as a means to increase the number of target molecules, but it is tedious and time-consuming. For a description of methods to clone nucleic acid fragments see Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual. A linear amplification of the target nucleic acid sequence using multiple rounds of primer extension; Proudfoot et al., Science 209:1329-1336, 1980; or the exponential amplification as described in detail in Mullis, et al., U.S. Pat. No. 4,683,195 and Mullis, U.S. Pat. No. 4,683,202 may also be utilized. It is necessary to either remove or inactivate the unincorporated nucleotides and primers of the amplification reaction. Non-labeled, unincorporated nucleotides will allow primer extension to occur beyond the nucleotide of

interest. Similarly, amplification primers that are free in solution can be extended and provide incorporation at positions other than the position of interest. Various methods obvious to those skilled in the art of molecular biology are available for removing unincorporated nucleotides and primers. However, since we desire a method that is rapid and automatable, the preferred form of separation is one utilizing attachment of the amplified nucleic acid product to a solid support with subsequent washing steps. An avidin-biotin system is preferred.

The template may be RNA or DNA, and may be double or single stranded. If double stranded, it is necessary to denature the strands to allow hybridization between the template strand and the oligonucleotide probe. Methods for this denaturation and subsequent hybridization step are well known to those skilled in the art of sequencing. However, since it is well known that formation of the hybrid between the oligonucleotide probe and the nucleic acid strand containing the target nucleic acid sequence can be inhibited by the complementary, non-template strand, the preferred method is to physically separate the template and the non-template strand after a denaturation step. The template strand can either be the strand present on the solid support (Figure 2), or a strand that is free in solution (Figure 3).

By design of an appropriate probe and utilizing the appropriate nucleic acid strand as template, essentially all nucleotide positions, even those at the end of a linear nucleic acid molecule can be assayed (Figure 4). In practice, since amplification from a complex nucleic acid mixture will at times give several different amplification products, the preferred method is to

utilize an oligonucleotide probe that is different from the primer(s) used in the amplification process.

The probe that has formed a duplex (hybrid) within the template is then subjected to enzymatic primer extension with enzyme such as primer-dependent DNA Polymerases and Viral Reverse Transcriptases, including AMV Reverse Transcriptase, various eukaryotic primer-dependent DNA Polymerase and DNA Polymerase I from E. coli (Klenow fragment). The basic elements required for execution of primer extension reactions are reviewed in Mullis et al., U.S. Pat. No. 4,683,195 and Mullis, U.S. Pat. No. 4,683,202, and include definition of a primer, size of primers, preparation of oligonucleotide primers, methods for separating strands of double stranded nucleic acid, preferable ratio of primer to template, conditions for mixing and annealing primer to template strand, and conditions for extending the primer in a 5' to 3' direction.

The enzyme used in the primer extension reaction should not exhibit exonuclease activity upon the components of the reaction. For example, either the enzyme should be free of 3' to 5' exonuclease activity or the probe should be of such composition as to resist such a degradation activity. Examples of this patent were performed under the former condition.

Adaptations of and alternatives to the primer extension technique can also be used with the process of the present invention. Double stranded nucleic acid targets can be used to generate both the template and primer strands, thereby eliminating the primer-template annealing step. By enzymatic or chemical treatment of the double stranded nucleic acid, molecules can be produced that have a recessed 3' strand and an overhanging 5' strand and thus are substrates for

nucleotide addition by a DNA polymerase. For example, cleavage of DNA with many restriction enzymes generates 5' overhangs that are substrates for DNA polymerases. Also, there are 3' exonucleases that remove 3' nucleotides from double-stranded DNA, producing molecules with 3' recessed strands and 5' overhanging strands.

The hybridizing and extending steps can be performed in solution or in solid phase reactions. The detection can also be in solution, after attachment to a solid phase, or after passing through a gel such as acrylamide or agarose. However, as previously mentioned, the first two methods are preferred for they avoid the time-consuming gel assay. Without a gel assay, it is necessary to separate the unincorporated labeled chain-terminators after the elongation step. Note that in the present invention, it is not necessary to wash away the excess oligonucleotide probe that did not hybridize, since the unextended probe does not contain a label.

There are four general forms of such separation: (1) immobilizing the elongated probe or hybrid selectively (e.g. by attaching to a binding segment on the analyte strand or on the probe) and separating away unincorporated, labeled nucleotide substrate together with sample polynucleotides that probe did not bind to; (2) immobilizing the elongated probe or hybrid non-selectively with other polynucleotides and separating away the unincorporated, labeled nucleotide substrate; (3) separating the unincorporated, labeled nucleotide substrate without immobilizing the elongated probe or hybrid, and (4) inactivating the label associated with unincorporated nucleotide substrate such that it is no longer detectable by the assay method employed.

Form (1) is the preferred method for it offers improved specificity and signal concentration in that a binding group can be captured specifically by a solid phase material. For example, a pendant biotin or biotin-dUTP incorporated into the probe can be specifically captured by avidin-coated materials such as avidin-agarose, avidin-coated magnetic beads, or avidin-coated microtiter wells. Another example might be the use of an oligonucleotide probe with a 5'-extension that is nonhomologous to the target sequence. This portion of the probe can then be used to capture the elongated probe (or hybrid) to a solid support that contains the complementary sequence. In such case, elongated probe or hybrid can be captured specifically and in high concentration on the solid phase, with the major other material captured (unhybridized probe) not causing non-specific signal.

If the capture onto solid support is due to a binding system present on the probe, then it should be apparent that any label generated in the assay from primers other than the probe would not be captured and therefore absent when analysis is performed.

Form (3) is also superior to many conventional probe assays where the probe is labeled before the elongation step, since separating a labeled oligonucleotide probe from a labeled, but short primer-elongation product is more difficult than separating the same labeled, primer-elongation product from the unincorporated, labeled nucleotide substrates.

Once the elongated probe is isolated from unincorporated, labeled nucleotide substrate, detection can proceed in a conventional fashion, either on the solid phase or otherwise. It should be apparent that the binding system used in forms (1) and (2) of the

present method should be independant of the binding system used to attach detectable label to the modified nucleotides during the detection step.

Crucial to this invention, are the chain-terminating, detectably labeled nucleotide substrates. Detectably labeled does not mean that the detectable signal must be present at the time of incorporation. The fluorescent substrates described below require activation. Detectably labeled does not necessarily mean that the nucleotide substrates carry a reporter such that there is not only the ability to detect the label, but also to identify the nucleotide. If only one nucleotide is present in the reaction, then detection of incorporation is sufficient for identification. The modified dideoxy-nucleotide substrates described in Prober et. al (EP-A 252683) or the DyeDeoxy terminators (a trademark of Applied Biosystems, Inc., Foster City, California) are examples of chain-terminating detectably labeled nucleotide substrates. However, unlike sequencing using fluorescently labeled chain-terminating nucleotides, there is essentially no requirement in this method that each of the modified nucleotides have a similar mobility shift when run on a sequencing gel. In the preferred embodiment, four chain-terminating nucleotides that are distinguishably labeled are present in each reaction. The need for four different labels is eliminated if the number of reactions per sample are increased (Figure 6). Unlike reports in the prior art, all four chain-terminating nucleotides may be present in the initial reaction, but only one must be detectably labeled. Unlike nucleic acid sequencing using chain terminators, the chain-elongating dNTP substrates are not a component of the reaction of the present invention.

The chain-terminating nucleotides described in the present invention are labeled with a fluorescent signal generator (reporter). A suitable fluorescent reporter is one that can be detected in its unprotected form at or below the level of detection that can be quickly achieved with ^{32}P, i.e., about 10^{-14} moles. Specific desirable characteristics may include a large coefficient of extinction in the region of excitation, a high quantum yield, an optimal excitation or emission wavelength (preferably above 350 nm), and photostability. For example, fluorescent dyes that are efficiently excited by an argon laser are desirable because of the low cost of this laser.

Preferably in its unprotected form, the reporter is a fluroescent dye chosen from the group consisting of xanthenes (e.g., fluoresceins, eosins, erythrosins), rhodamines (e.g., tetramethylrhodamine, Texas Red®), benzamidizoles, ethidiums, propidiums, anthracyclines, mithramycins, acridines, actinomycins, merocyanines, coumarins (e.g., 4-methyl-7-methoxycoumarin), pyrenes, chrysenes, stilbenes, anthracenes, naphthalenes (e.g., dansyl, 5-dimethylamino-1-naphthalenesulfonyl), salicyclic acids, benz-2-oxa-1-diazoles (also known as benzofurazans) (e.g., 4-amino-7-nitrobenz-2-oxa-1,3-diazole), and fluorescamine. Useful forms of many of these dyes are commercially available. For a review of fluorescent dyes used in tagging DNA, see A. S. Waggoner, Chapter 1, Applications of Fluorescence in the Biomedical Sciences, ed. by D. L. Taylor, et al., Alan R. Liss, New York (1986). An extensive description of chain terminator labeling is found in U.S. Application No. 07/057,566 filed June 12, 1987, incorporated herein by reference.

The present invention is further illustrated by reference to Figures 1-6.

In Figure 1a, an analyte strand (An) contains a nucleotide position of interest (N), the identity of which is to be determined by the assay, is defined as the first base of the analyte nucleic acid strand which is beyond the 5' end of the target nucleotide sequence in the 3' to 5' direction. A probe polynucleotide is produced as a reagent having a binding region complementary to the target nucleotide sequence (TNS). In this particular embodiment, the probe polynucleotide consists only of that complementary sequence; in other embodiments, the probe is extended in the 5' direction in a manner that does not interfere with the recognition and complementary base pairing to the target nucleotide sequence. The diagram in Figure 1a illustrates the double stranded nucleic acid region which forms when the probe binds to analyte strand An by complementary base pairing to the target nucleotide sequence TNS.

By contacting the double stranded region shown in Figure 1a with a DNA polymerase specific therefore, the 3' end of the probe will be utilized as a primer and elongated opposite the analyte strand An which serves as a template for nucleotide incorporation. As illustrated in Figure 1b, the nucleotide incorporated (N*) will be complementary to the nucleotide position of interest (N). In all illustrations, the * symbol is used to illustrate a detectable label attached to the nucleotide. The enzyme, primer and nucleic acid analyte are chosen together such that a nucleotide complementary to the target nucleotide of interest is incorporated. For example, if analyte strand An is DNA, then a reverse transcriptase, a primer dependent prokaryotic DNA polymerase (e.g. the Klenow fragment of E. coli DNA

Polymerase I or TAQ polymerase) or a eukaryotic DNA polymerase may be used, with the probe being DNA or RNA..

Figures 1c-e are examples to illustrate the more schematic drawings of Figures 1a&b. In Figure 1c, the target nucleotide of interest is a T. After chain elongation of the primer with a DNA polymerase in the presence of detectably labeled chain terminators, the complementary nucleotide, A*, will be covalently attached to the primer (Figure 1d). If the nucleotide of interest was a C, then the complementary nucleotide that is incorporated will be a G* (Figure 1e). If the nucleic acid sample being analyzed contains molecules of several types, then several different nucleotides may be incorporated and covalently attached to the primer (e.g. if both T and C are present in the position of interest in a portion of the molecules, then a result corresponding to Figures 1d&e may occur within one sample being analyzed).

Figures 1f-h are very similar to Figures 1c-e except that the opposite nucleic acid strand is utilized as template thus the oligonucleotide probe is chosen to correspond to a different TNS.

Figures 2 & 3 illustrate in schematic form the sequence of events that comprise preferred embodiments for carrying out the present invention. In Figure 2 the immobilized strand is used as the template (An), and in Figure 3 the eluted or non-immobilized strand is used as template. As shown, the polymerase chain reaction (PCR) may be used to generate the template strand in increased quantity before analysis. The removal of the unincorporated nucleotides and primers is essential, and can be performed by binding the double stranded PCR product to a solid support, e.g. by a biotin (B) -

streptavidin complex, and rinsing away the unbound material. The two strands are then denatured (e.g. by addition of NaOH) and only the template strand is retained for the reaction. In Figure 2 the immobilized template strand is rinsed, while in Figure 3 the soluble, eluted strand is used as template after neutralizing the NaOH solution. The probe oligonucleotide is then hybridized to the template strand and the hybridized probe is elongated by addition of a single, chain terminating nucleotide. The enzyme utilized in the reaction is a DNA polymerase such as reverse transcriptase and all four chain terminating nucleotides may be present, although only one must be detectably labeled. The unincorporated nucleotides are removed from the reaction by washing. Note that in Figure 3, the template strand was not previously immobilized, so the probe oligonucleotide can now be captured onto solid support for efficient washing. The nature of the label present on the elongated primer may be measured directly after efficient removal of the unincorporated substrate. That is, the primer may still be bound to the solid support, either directly as shown in Figure 3 or indirectly through the hybrid formed with the analyte strand (Figure 2 without the final denaturation step). For the particular brand of streptavidin coated magnetic beads used in our examples, the labeled primer is released from the beads after heating in the presence of formamide and EDTA. The magnetic beads do not interfere with standard gel electrophoresis although they are loaded into the sample well along with the sample. If the sample is assayed through a capillary, then the beads may obstruct flow and should be removed.

Figure 4 illustrates that the nucleotide of interest (N) can even be located at the end of a nucleic acid strand. This is different from that of multiple nucleotides during the reaction.

Figure 5 illustrates that the assay is useful for detecting insertions and deletions as well as the point mutations illustrated in Figures 1c-h. In Figure 5a, two template nucleic acid strands are drawn, with the upper strand differing from the lower strand by the presence of two T's (note that one strand may be considered to have a deletion, or the other strand may be considered to have an insertion). Figures 5b&c illustrate one possible choice of probe and the resulting difference in the nucleotide incorporated when the two different strands are used as the template, An.

Figure 6 illustrates how the number of assays must increase if all four chain terminating nucleotides are not detectably labeled and distinguishable one from another.

The number of reactions required to identify the nucleotide of interest in a given sample is dependent upon how many of the different, possible substrates are detectably labeled and distinguishable.

Method 1 (Preferred Method)

All possible nucleotide substrates are present in reaction and all four are detectably labeled and distinguishable from each other (e.g., dd**A**TP*, dd**T**TP*, dd**C**TP*, dd**G**TP* are provided as substrate)

One Reaction:

Result: Only dd**A**TP* is incorporated and detected.

Conclusion: Since the other nucleotides were not detected, only **T** is present at the nucleotide of interest. No other reactions are required.

Method 2

All possible nucleotide substrates are present in each reaction but perhaps only 2 can be detectably labeled such that they are distinguishable from each other.

1st Reaction: provide substrates dd**A**TP*, dd**G**TP*, dd**C**TP, dd**T**TP.

Result: Only dd**A**TP* is incorporated and detected.

Conclusion: **T** is present at the nucleotide position, and **C** is not present. But do not know if either G or A are present.

2nd Reaction: provide substrates dd**C**TP*, dd**T**TP*, dd**A**TP, dd**G**TP.

Result: There was no detectable incorporation.

Conclusion: **G** or **A** are not present at the position of interest.

(To provide evidence to support this further, one could use the same substrate mixtures except monitor the incorporation using the complementary strand as the anlayte and the different, but appropriately positioned primer.)

Method 3

All possible nucleotide substrates are present in each reaction but perhaps only one label is available for substrate labeling (e.g., the same as when radioactively labeled ddNTP's are utilized).

1st Reaction: provide substrates dd**A**TP*, dd**G**TP, dd**C**TP, dd**T**TP.

Result: Only dd**A**TP* is incorporated and detected.

Conclusion: **T** is present at the nucleotide position, but after only this first reaction, it cannot be stated that another nucleotide is not also present.

2nd Reaction: provide substrates dd**A**TP*, dd**G**TP*, dd**C**TP, dd**T**TP.

Result: No detectable incorporation.

Conclusion: **C** is not present at the nucleotide position of interest.

3rd Reaction: provide substrates dd**A**TP, dd**G**TP, dd**C**TP, dd**T**TP.

Result: No detectable incorporation.

Conclusion: **G** is not at the nucleotide position of interest.

4th Reaction: provide substrates dd**A**TP, dd**G**TP, dd**C**TP, dd**T**TP*.

Result: No detectable incorporation.

Conclusion: **A** is not at the nucleotide position of interest.

EXAMPLES

The following examples are offered by way of illustration and are not intended to limit the invention in any manner.

EXAMPLE 1

Aims:

1. To demonstrate that the claimed method can be used to identify the nucleotide present at a defined position on a nucleic acid strand in a single reaction.

2. To illustrate that a single fluorescently labeled nucleotide can be incorporated in the assay using a commercially available enzyme preparation.

3. To illustrate that the unincorporated labeled substrate can be efficiently removed without time-consuming centrifugation or column chromatography.

4. To illustrate the use of a DNA strand labeled at the 5' end with Biotin and bound to a solid support as the analyte strand.

5. To illustrate the ability to distinguish between three DNA samples by incorporation of an A (wildtype allele), a G (mutant allele), or both A & G (heterozygote) as the complementary nucleotide opposite the nucleotide position of interest.

6. To illustrate that the fluorescent nucleotide substrates can be detected and distinguished on the Genesis 2000 DNA analysis unit either by gel electrophoresis or by passing the sample through a capillary.

Definition of the Nucleotide Position of Interest:

In this example, the nucleotide position of interest is that of the lower strand at nucleotide position 5430 of the mouse RNA polymerase II largest subunit gene as described by publication in the GenBank database, accession M12130 for the locus RO:Musrpolii2. A 602 nucleotide portion of this sequence from nucleotide 4915 to 5517 is illustrated in its double stranded form in Figure 8, with the nucleotide position of interest for this example being at position 5430 on the lower strand (occupied by a bold-faced T in the sequence of the Wildtype allele which is shown in this Figure 8).

Definition of the Target Nucleotide Sequence:

In this example, the target nucleotide sequence (TNS) is chosen as the 21 nucleotide sequence (3'TAACGACAACAGCCCGTCGTC5') that immediately flanks the nucleotide of interest such that the nucleotide position of interest is the next contiguous nucleotide in the 3' to 5' direction on that nucleic acid strand (see Figure 1c).

The oligonucleotide probe:

In this example, the oligonucleotide probe consisted of the 21 nucleotide sequence 5' ATTGCTGTTGTCGGGCAGCAG 3' (probe A of Figures 8 and 9), and is perfectly complementary to the target nucleotide sequence defined above. It is synthesized on an Applied Biosystems DNA synthesizer and further purified by HPLC to consist of a single oligonucleotide species, 21 nucleotides in length. (methods as described in *Oligonucleotide Synthesis, A Practical Approach* ed. M.J. Gait, IRL Press 1984).

Starting biological sample:

In this example, the claimed method will be illustrated using three different amplified DNAs in three separate, but similar reactions. The three starting biological materials used for the amplification process are each known to contain the mouse RNA polymerase II gene. The three samples are designated Wildtype, Mutant, and Heterozygote. They are known to differ at nucleotide position 5430 (as shown in bold faced type in Figure 9), with the Wildtype allele containing an A-T base pair, the Mutant allele containing a G-C base pair, and the Heterozygous sample containing an equal mixture of these two alleles. The

starting biological materials are obtained from J. Corden and are as described in Bartolomei and Corden, Molec. and Cell. Biol. 7:586-594, 1987. The Wildtype and Mutant alleles are provided as bacterial strains containing the recombinant plasmids pE26-4 and pE26-7 respectively. The biological sample designated as Heterozygous is obtained as a cell line A21. DNA of each of the recombinant plasmids is prepared by standard molecular biology procedures (described in Sambrook et al., *Molecular Cloning: A Laboratory Manual* 1989), and genomic DNA is prepared from the A21 cell line as described in Corsaro and Pearson, Somatic Cell Genet. 7:603-616, 1981.

Amplifying a segment of DNA containing the nucleotide position of interest and the target nucleotide sequence:

As shown in Figure 8, the target nucleotide sequence and the nucleotide position of interest are within a 602 base pair segment of the RNA polymerase II gene. The copy number of this segment is increased using exponential amplification, using DNA of each of the three biological starting materials described above. The oligonucleotide primers used for PCR amplification of the region of interest in the RNA polymerase II gene are designated PCR amplification Primer 1 (5' CAGACATTTGAGAATCAAGTGAATCG 3') and PCR amplification Primer 2 (5' BCTCGGCTCTCAGGACCATAATCAT 3') where B=biotin (see Figure 8). They are synthesized by standard phosphoramidate chemistry on an Applied Biosystems DNA synthesizer. For the biotinylated primer, the biotin moiety is added at the 5' end during synthesis as described in Cocuzza US patent 4,908,453. All such oligonucleotides used in this patent are prepared for the inventors by the Du Pont

Oligonucleotide Synthesis Facility under the direction of C. Burns, although commercial firms are available for attaining such items.

For PCR amplification, thirty-three picomoles of each primer (one primer biotinylated at the 5' end) is mixed with approximately 1 ug of genomic DNA (or 0.1 ug of plasmid DNA) in a 50 µl reaction mixture containing 60 mM KCl, 15 mM Tris (pH8.8), 2.75 mM MgCl2, and dATP, dGTP, dCTP, and TTP each at 200 uM. The mixture is incubated at 95°C for 2 min to separate the DNA strands and cooled on ice; 2.5 units of TAQ polymerase (AmpliTaq, Perkin Elmer Cetus) is added, and the reaction mixture is overlaid with approximately 35 µl of mineral oil. The amplification conditions varied slightly in the course of the experiments, but are usually performed in a Perkin-Elmer/Cetus thermal cycler using an initial cycle consisting of 4 min 94°C, 45 sec 55°C, 5 min 68°C, followed by 35 cycles with the same parameters except the denaturation at 94°C is 1 min. Following PCR amplification, 10-15 µl aliquots of the amplified fragment are run on a 1.5% agarose gel and visualized by ethidium bromide staining using standard procedures to ensure that an amplified fragment of the expected size is produced. Utilizing these primers and any of the three DNAs described above, a 602 bp PCR amplification product of double stranded DNA is consistently obtained. Aliquots of the remainder of the PCR amplification sample are then utilized in the method of this invention.

Preparation of the analyte strand:

The removal of unincorporated nucleotides and any interfering primers before performing the method of this invention is essential. In this example, the double-

stranded PCR amplification product contains a biotin moiety due to the biotin originally presnet on PCR amplification Primer 2. Thus, the separation is done by binding the biotinylated PCR amplification product to a streptavidin-coated solid support and rinsing away the non-biotinylated, PCR amplification Primer 1 and the unincorporated nucleotides. The support-bound PCR amplification product is then denatured using NaOH, the complementary, non-analyte strand is removed and the remaining analyte strand which is still bound to the solid support is rinsed and ready for the primer extension reaction. These steps are illustrated in the schematic drawing of Figure 2 and described below as steps 1-6.

1. Magnetic, streptavidin-coated beads from the Dynal corporation (Dynabeads M-280 Streptavidin, at 6 x 108 beads/ml) are washed and resuspended at the same concentration in Triton Wash Solution [0.17% (w/v) Triton X-100, 100 mM NaCl, 10 mM Tris-HCl pH7.5, 1 mM EDTA] essentially as described in the Application Brief 25 for the Genesis 2000 DNA analysis system.

2. Approximately 20 µl of double-stranded DNA template, amplified using PCR amplification with one of the two primers labeled with biotin, are mixed with 20 µl of washed Dynabeads and incubated at 37°C for 30 minutes. This mixture is gently shaken intermitently in order to keep the magnetic beads in solution.

3. After this incubation, the tube containing magnetic beads and DNA is placed near a magnet to draw the beads to one side of the tube. After approximately four minutes of magnetization, the supernatant is removed.

4. The beads (with DNA bound) are then washed three times with TE buffer (10 mM Tris pH8, 1 mM EDTA) using

magnetization for removal of the supernatant which contains dNTP's and non-biotinylated PCR amplification primer. Care is taken that the beads did not dry between washes.

5. After the final wash, 16 µl of sterile distilled water is added to the bead-bound DNA.

6. The double-stranded DNA is denatured by addition of 4 µl of 0.5M NaOH, 2 mM EDTA solution, and incubated at room temperature for 5 min. Afterwards, the sample is magnetized and the supernatant removed. (The supernatant may be kept if the non-bead bound strand is to be used as template - e.g. Example 5). The bead-bound DNA pellet is gently resuspended in 100 µl TE buffer to neutralize any NaOH remaining.

Formation of the analyte-probe hybrid and enzymatic extension of the probe with a chain terminating nucleotide complementary to the nucleotide position of interest:

7. To use the bead-bound DNA as template, the TE buffer is removed following magnetization, and 7 µl of the following solution is added:

2 µl sterile water
1 µl 125 µM ddTTP (unlabeled)
1 µl 125 µM ddCTP (unlabeled)
1 µl 6.6 µM oligonucleotide probe
2 µl 5X RT buffer [supplied by Invitrogen for use with Reverse Transcriptase]

8. Incubate at 50°C for 5 min, then transfer to 37°C for an additional 7 minutes, and then transfer to ice.

9. The fluorescently-labeled chain terminators and the primer-dependent enzyme are then added to the

reaction. In this particular Example 2, it is as follows:

1 µl of 125 µM SF-ddGTP-505
1 µl of 125 µM SF-ddATP-512
0.5 µl Invitrogen Reverse Transcriptase (10µ/µl)

10. The labeling reaction is at 42°C for 10 minutes, and then the reaction is placed on ice and 100 µl of TE is added.

11. The sample is again magnetized for 4 minutes and the supernatant removed, followed by 3 washes of 100 µl TE buffer (magnetization between each wash) to remove unincorporated nucleotides.

12. The final supernatant is removed and the magnetic beads (with DNA bound) are resuspended in 6 µl FE solution (95% formamide, 25mM EDTA) and stored -4°C until further use.

Detection of the chain terminating nucleotide attached to the probe:

13. The sample from step 12 is diluted 1:16 fold further in FE containing crystal violet, for easier visualization in loading the sample and to get the sample in a reasonable concentration for detection by slab gel electrophoresis on the Genesis 2000 DNA analysis system (methods as described by the instrument documentation, with a few parameters described in more detail below).

14. "Lane Finding" for the Genesis detection system is performed manually using a primer fluorescently labeled at the 3' end prepared in advance using terminal transferase and a fluorescent ddNTP as substrate as described in Trainor and Jensen, Nucl. Acids Res. 16:11846, (1988) that is electrophoresed into each

sample lane approximately thirty minutes prior to briefly pausing the machine and then loading the reaction samples into the lanes. Finding lanes in this manner allowed the recording of the fluorescence detection in each lane to begin early enough such that any unincorporated labeled substrate remaining in the reaction would be recorded.

15. Just prior to loading 1 µl of the reaction sample onto the gel for electrophoresis and subsequent detection, it is combined with approximately 1 µl of a shorter, fluorescently-labeled control primer, and heated 95°C for 2 minutes. (This control primer is added as a mobility standard, but is later found to be an unnecessary component and is omitted in later electrophoresis runs).

The first three lanes shown in Figure 10a are results from running such samples from each of the three reactions. The position of the peak corresponding to the elongated probe is designated WT, Mutant, and Het for the three reactions of this example. In these examples, the peak at position S corresponds to the control primer that is added at the time of electrophoresis. As can be noted, there is very little unincorporated fluorescent nucleotide (peak at position U) remaining in each of the three reactions due to the washes of step 11.

Determining the identity of the chain terminating nucleotide that is added to the probe:

The fluorescently labeled chain terminators used in the examples of this patent are either purchased from duPont NEN Biotechnology Systems (Boston, MA), or obtained as a kind gift from Dr. Douglas Amorese of that firm. The nature and detection of these SF-ddNTPs are

described in Prober et al. (1988) *Science* **238**, 336-341. In brief, the chain terminators are distinguished by a ratio of the measured fluorescence from two photomultiplier tubes (PMT). Each PMT value is displayed as either a red or green mark on the output computer monitor, with a sample forming a peak as it passes by the excitory laser. (In the black and white Figures required for this patent application, the original color of each line of the sample peak is indicated). Unlike the normal method of fluorescent base detection when multiple peaks of a sequencing reaction are being analyzed, the commercial Genesis 2000 software is unable to determine the identity of the fluorescent nucleotides (base call) in this application, for only a few peaks are present in the lane. It is therefore necessary to prepare a one time calibration on the instrument by preparing a set of expected values for the fluorescently labeled chain terminators at various dilutions in FE (95% formamide, 25mM EDTA). It is important to note that the commercial instrument is designed to have a non-linear response of the two PMTs when the voltage is too high. We experimentally determined that the ratio obtained for the green peak height to the red peak height for a given fluroescent substrate is relatively invariant from experiment to experiment over the range of 0.1-9 volts. Thus the green to red ratio of a peak is only determined if the reaction samples are within this voltage range.

The result of such a calibration (+/- one standard deviation) is shown in Figure 7a by two different assay methods. The samples are either electrophoresed on a urea-polyacrylamide gel by standard gel electrophoresis procedures for the Genesis 2000, or syringe-loaded into a single, empty capillary (Part # TSP530700 from

Polymicro Technologies, Phoenix, AZ) positioned in front of the excitatory laser beam on a Genesis 2000 unit. A photo and full description of this modified Genesis apparatus is given in Zagursky and McCormick, Biotechniques 9:74-79 (1990). (Care should be taken during capillary alignment to avoid electrical shock or direct eye contact with the laser beam.) An example of the type of data collected using standard gel electrophoresis methods on the Genesis is shown in Figure 7b for SF-ddGTP-505 and SF-ddTTP-526. Figure 7c illustrates the type of measurement made when the fluorescent substrate SF-ddGTP-505 or SF-ddCTP-519 is loaded via syringe into the capillary mounted onto the Genesis 2000 detection system. The multiple peaks represent the same sample being pushed several times in front of the laser beam. Although the green/red ratio for a particular fluorescent substrate is different from that of Figure 7b, the nucleotides can be distinguished in this new detection system at concentrations similar to that of gel electrophoresis as illustrated in Figure 7a. (However, since the rate at which the sample is manually pushed in front of the detection system is not uniform, the overall peak height is not reproducible in this experiment).

Regardless of whether the sample is electrophoresed through a gel matrix or pushed through a capillary, the SF-ddNTPs are distinguishable (Figure 7a).

16. To determine the identity of the chain terminating nucleotide attached to the probe, the PMT ratio (green/red peak height) is measured for each sample using the position of the initial rise of the peak as the baseline.

For accurate determination of this ratio, the three reaction samples of this example (i.e. the WT, Mutant,

and Het peak) are rerun at lower dilution (since the voltage of two of them are originally too high as shown in Figure 10a). The resulting sample peaks are displayed in Figure 10b with a smaller display window for easier measurement. In this example, the measured green/red ratios are as follows:

WT = 1.55 Mutant = 2.5 HET = 1.9

A comparison of these values to the calibration shown in Figure 7a illustrates that these ratios correspond to the expected incorporation of SF-ddATP-512 when the Wildtype allele is the source of the analyte strand, incorporation of SF-ddGTP-505 when the Mutant allele is the source of the analyte strand, and a mixture of both fluorescent nucleotides in approximately equal proportions when the analyte strand is derived from a heterozygous source (i.e. approximately equal number of Wildtype and Mutant analyte strands).

Identifying the nucleotide of interest as the nucleotide complementary to the chain terminating nucleotide which is added:

17. The nucleotide at the position of interest is the nucleotide complementary to the nucleotide that is incorporated.

Therefore, the conclusion for the three samples of this example are as expected:

The reaction performed on the Wildtype allele indicates that it does contain a thymidine (T) on the lower strand at nucleotide position 5430 of the mouse RNA polymerase II largest subunit gene, for the nucleotide incorporated is SF-ddATP-512. The reaction performed on the Mutant allele indicates that it does

contain a cytosine (C) on the lower strand at nucleotide position 5430 of the mouse RNA polymerase II largest subunit gene, for the nucleotide incorporated is SF-ddGTP-505. The reaction performed on DNA originating from the A21 cell line (which is known to be heterozygous for the wildtype and mutant allele) contains an equal number of thymidine (T) and cytosine (C) residues at the position of interest, for an approximately equal number of SF-ddATP-512 and SF-ddGTP-505 are incorporated onto the probe.

EXAMPLE 2

Aims:

1. To illustrate the ability to incorporate and distinguish a fluorescently labeled SF-ddCTP-519 and SF-ddTTP-526 in the practice of this invention.

2. To illustrate the ability to use the complementary strand to that used in Example 1 as the bead-bound analyte strand and to illustrate the use of another oligonucleotide sequence as the probe (probe B of Figure 8).

Definition of the Nucleotide Position of Interest:

In this example, the nucleotide position of interest is that of the upper strand at nucleotide position 5430 of the mouse RNA polymerase II largest subunit gene as described by publication in the GenBank database, accession M12130 for the locus RO:Musrpolii2. A 602 nucleotide portion of this sequence from nucleotide 4915 to 5517 is illustrated in its double stranded form in Figure 8, with the nucleotide position of interest for this example being at position 5430 on the upper strand (occupied by a bold-faceted A in the sequence of the Wildtype allele which is shown in this Figure 8).

Definition of the Target Nucleotide Sequence:

In this example, the target nucleotide sequence (TNS) is chosen as the 21 nucleotide sequence (5'ATGTAGAGGGCAAGCGGATCC3') that immediately flanks the nucleotide of interest such that the nucleotide position of interest is the next contiguous nucleotide in the 3' to 5' direction on that nucleic acid strand (see also Figure 1f).

The oligonucleotide probe:

In this example, the oligonucleotide probe consisted of the 21 nucleotide sequence 5' GGATCCGCTTGCCCTCTACAT 3' (probe B of Figures 8 and 9), and is perfectly complementary to the target nucleotide sequence defined above. Synthesis and purification is as described in Example 1.

Starting biological sample:

In this example, the claimed method will be illustrated using two of the same starting biological samples as described in Example 1: that of the Wildtype and Mutant. They are prepared as described in Example 1.

Amplifying a segment of DNA containing the nucleotide position of interest and the target nucleotide sequence:

The region of interest is amplified from the Wildtype and Mutant samples using methods as described in Example 1 with PCR amplification Primer 1 and PCR amplification Primer 2, except in this Example 2, the PCR amplification Primer 1 is biotinylated at the 5' end and the PCR amplification Primer 2 is not.

Preparation of the analyte strand:

The methods are as described in Example 1 steps 1-6, however in this Example, it is the complementary strand which is bound to the solid support, for this strand contains the biotin from the PCR amplification reaction.

The invention is practiced as in the steps of Example 1 on the Wildtype and Mutant analyte strands with the following exceptions:

a) The two unlabeled nucleotide substrates in step 7 are ddGTP and ddATP .

b) The two fluorescently labeled nucleotide substrates in step 9 are 1 µl of 30 uM SF-ddCTP-519 and 1 µl of 125 µM SF-ddTTP-526.

The results shown in Figure 11 illustrate that the Wildtype and Mutant allele have green/red ratios of 0.35 and 0.7 respectively when probe B is used. The calibration graph of Figure 7a shows that this corresponds to the incorporation of SF-ddTTP-526 and SF-ddCTP-519 for the Wildtype and Mutant allele.

THE CONCLUSION for the two samples of this example are as expected (see Figure 9):
The reaction performed on the Wildtype allele indicates that it does contain a adenine (A) on the upper strand at nucleotide position 5430 of the mouse RNA polymerase II largest subunit gene, for the nucleotide incorporated is SF-ddTTP-526. The reaction performed on the Mutant allele indicates that it does contain a guanine (G) on the upper strand at nucleotide position 5430 of the mouse RNA polymerase II largest subunit gene, for the nucleotide incorporated is SF-ddCTP-519.

EXAMPLE 3

Aim:

1. To illustrate that under the conditions of Examples 1 and 2, that in some cases, the wrong nucleotide will be incorporated if the correct nucleotide is missing from the reaction (i.e. the problem with many of the assays discussed in the prior art is that of significant misincorporation in reactions where the correct nucleotide is not provided.

For this example, the same two reactions as described in Example 2 are performed with the following exceptions:

a) For the reaction with the Wildtype allele in step 9 the SF-ddTTP-526 is omitted and the only fluorescent substrate in the reaction is 1 µl of 125 µM SF-ddCTP-519. (Unlabeled ddTTP is also absent in the reaction).

b) For the reaction with the Mutant allele in step 9 the SF-ddCTP-519 is omitted and the only fluorescent substrate in the reaction is 1 µl of 125 µM SF-ddTTP-526. (Unlabeled ddCTP is also absent in the reaction).

The results shown in Figure 12 illustrate that for the Wildtype allele (upper panel), there is no significant misincorporation of SF-ddCTP-519 as a complementary base for the adenine (A) present as template for the primer extension. This is true although the SF-ddCTP-519 is present at a higher concentration than in Example 2. No conclusions can be made with respect to whether A or G is incorporated, for these nucleotides although present, are not fluorescently labeled in the reaction. The lower panel of Figure 12 illustrates a significant level of misincorporation of SF-ddTTP-526 as a complementary base

for the guanine (G) present on the Mutant analyte strand for the primer extension of probe B (refer to Figure 9).

These results suggest that in some cases (e.g. as in the upper panel of Figure 12) there will be very little misincorporation when the correct, complementary base is not provided in the reaction. Therefore, in such cases, the method of the present invention can be performed without all four bases present. However, the lower panel of Figure 12 clearly suggests that in other instances, the perferred form of this invention would be to perform the reaction with all four ddNTPs present.

EXAMPLE 4

Aims:

1. To illustrate that the use of all four detectably labeled nucleotides in a single reaction gives essentially the same PMT green/red ratio as compared to Example 2 when the correct base is present.

2. To illustrate that a lowering of the concentration of ddNTP substrate by 5 fold may improve the ability to rinse away unincorporated ddNTP's without affecting the ability to measure the sample.

Sample Preparation

The reaction is the same as that described for Example 2 except the unlabeled nucleotides are omitted from step 7 and all four ddNTPs are present in step 9 at 1/5 the concentration. The results shown for Mutant and Wildtype amplified alleles are shown in the two panels of Figure 12. The Mutant template resulted in a green/red ratio = 0.77 suggesting the correct incorporation of SF-ddCTP-519, while the Wildtype template gave green/red = 0.42 suggesting correct incorporation of a SF-ddTTP-526 (again see

discriminatory values given in Figure 7b). In these lanes, as in other lanes of the experiment, there is essentially no peak of unincorporated nucleotides present in the sample.

EXAMPLE 5

Aims:

1. To illustrate the use of a nucleic acid strand that is not bound to a solid support as the analyte strand.

2. To practice the method of this invention on a totally different biological sample than that used in Examples 1-4.

Definition of the Nucleotide Position of Interest:

In this example, the nucleotide position of interest is that of the lower strand, occupied by a circled G on Figure 14. The nucleotide sequence shown is a portion of the Wildtype A1 gene of maize (Schwarz-Sommer et al., EMBO J. 6:287-294 (1987).

Definition of the Target Nucleotide Sequence:

In this example, the target nucleotide sequence (TNS) is chosen as the 21 nucleotide sequence (3'GACGAACTCCTAGCTCATCAC5') that immediately flanks the nucleotide of interest such that the nucleotide position of interest is the next contiguous nucleotide in the 3' to 5' direction on that nucleic acid strand.

The oligonucleotide probe:

In this example, the oligonucleotide probe consists of the 21 nucleotide sequence 5' CTGCTTGAGGATCGAGTAGTG 3' (Primer C of Figure 14), and is perfectly complementary to the target nucleotide sequence defined

above. Primer C is biotinylated at the 5' end during primer synthesis and is HPLC purified by methods described in Example 1.

Biological sample:

DNA sequence analysis of several alleles of the maize A1 gene had shown that among other differences, the a-dt mutant allele contains a G-C base pair while the Wildtype A1 allele had a C-G base pair at the nucleotide positions circled in Figure 14. For this example, total genomic plant DNA is prepared (method as described in Shepherd et al., Mol. Gen. Genet. 188:266-271 (1982) from a maize plant known to be homozygous for the a-dt mutant allele of the A1 gene.

Amplifying a segment of DNA containing the nucleotide position of interest and the target nucleotide sequence:

PCR amplification primers A & B homologous to a section of the maize A1 gene are designed, synthesized, and used to amplify the genomic fragment from total maize DNA containing the mutant a-dt allele of maize by methods described in Example 1. In this example, the Primer B is the PCR amplification primer containing biotin at the 5' end.

Preparation of the analyte strand is performed as in steps 1-6 of Example 1, except that the non-bead bound strand from step 6 will be used as the analyte strand below.

The sequence of events for using the non-bead bound strand as the analyte strand is illustrated in Figure 3 and the steps are given below.

7. After denaturation of the two PCR amplification strands, the strand that is present in 16 µl of basic

NaOH solution is carefully neutralized by addition of a few microliters of 0.5M HCl, monitoring the pH of the solution using pH paper. To this neutralized DNA sample (vol. aprox. 23 µl), the following addition is made:

8.0 µl of 5X RT buffer (Invitrogen)
2.8 µl of 1% Triton X-10
0.5 µl of 1 mM ddATP (unlabeled)
0.5 µl of 1 mM ddTTP (unlabeled)
4.0 µl of 6.6 µM Biotinylated nested primer

8. The sample is incubated 95°C for 2 minutes, 37°C 10 minutes, and then placed on ice.

9. The following additions are made:
2 µl of 125 µM SF-ddCTP-519
2 µl of 125 µM SF-ddGTP-505
1 µl of Reverse Transcriptase (Invitrogen 10 µ/µl)

10. The sample is incubated 42°C 10 minutes.

11. 15 µl of Dynabeads (prepared as in step 1) are added and followed by a 37°C incubation for 15 minutes with intermitent shaking. This is to promote binding of the nested, biotinylated primer (containing fluorescent label from the primer extension reaction).

13. The sample is magnetized for 4 minutes and unincorporated nucleotides present in the supernatant are removed.

14. The final bead pellet is washed 3 times with 100 µl TE (magnetizing each time to remove the buffer).

15. The final bead pellet is resuspended in 5 µl of FE (95% formamide, 25 mM EDTA).

16. 2 µl of this sample along with 1 µl of a smaller, control primer (Std) are heated for 2 min 95°C, before loading on a urea-polyacrylamide gel and electrophoresis on the Genesis 2000.

The panel of Figure 14 illustrates the peak due to incorpor ation of a base complementary to the nucleotide position of interest for the Mutant allele (a-dt). The green/red ratio is 2.6, consistent with the correct base G being added to the biotinylated, nested primer. Although this procedure clearly worked, correct neutralization of the template strand is a time-consuming process that is somewhat variable with respect to final pH and resulting salt concentration. The preferred method is therefore to work as in examples 4-6 with the biotinylated strand as template.

EXAMPLE 6

Aim:

1. To illustrate the use of another primer dependent DNA polymerase in practicing this invention.

2. To demonstrate that the method of this invention is capable of distinguishing a heterozygous DNA sample with one reaction (in this case, an equal number of cytosine and guanine nucleotides at the nucleotide position of interest).

This example is the same as Example 5 with the following exceptions:

a) Genomic DNA is prepared from maize leaf material as described in Example 5, but the exact nature of the nucleotide of interest in the sample is unknown until after the method of this invention is performed. (Standard DNA sequence analysis of this region of the DNA later confirmed that the biological material is heterozygous at the nucleotide position of interest with both cytosine (C) and guanine (G) being present in essentially equal amounts (data not shown).

b) In step 7, the following are the additions made to the neutralized, non-bead bound strand:

8.0 µl of 5X Sequenase buffer (200 mM Tris pH 7, 100 mM MgCl2, and 250 mM NaCl)
2.8 µl of 1% Triton X-100
0.5 µl of 1 mM ddATP (unlabeled)
0.5 µl of 1 mM ddTTP (unlabeled)
4.0 µl of 6.6 µM Biotinylated oligonucleotide probe (primer C)

c) In step 9, 1.5 µl of 100 mM Dithiothreitol is added (in addition to the fluorescent substrates SF-ddCTP-519 and SF-ddGTP-505), and 1 µl of 13 units/µl Sequenase Version II enzyme (a modified T7 DNA polymerase: US Biochemical Corp. US Patent 4,795,699) instead of the reverse transcriptase enzyme.

d) In step 10, the labeling reaction is at 37°C for 10 minutes.

The results of Example 6 are shown in Figure 16. It is seen that a single peak of incorporation appears, suggesting that the Sequenase II enzyme can also be used for the practice of this invention with no significant 3'-5' exonuclease activity. The green/red ration (=1.7) of this peak is as would be expected for a DNA sample that is heterozygous. That is, the calibration graph of Figure 7c indicates that the value of 1.7 is approximately equal distance between the expected values for incorporation of SF-ddCTP-519 (green/red approximately 0.9) and that for incorporation of SF-ddGTP-505 (green/red approximately 2.4). Note that in this example, SF-ddATP-512 that gives a green/red ratio of approximately 1.6 is not included in the reaction, thus the 1.7 ratio does not indicate the addition of an adenine.

In conclusion, in Example 6 it is determined that the nucleotide position of interest in the sample is

occupied by an approximately equal number of cytosine (C) and guanine (G) residues.

EXAMPLE 7

This example illustrates identification of a nucleotide of interest at a defined location in samples of a 700 bp DNA element, Cin1, from the analysis is carried out in a single reaction Northern Flint Line of Zea mays in a single reaction. The Cin1 element is repetitive in Zea mays.

The oligonucletodie probe is a 21 nucleotide sequence perfectly complementary to the target nucleotide sequence. The target nucleotide sequence is a 21 nucleotide sequence that immediately flanks the nucleotide of interest at position 500. Synthesis of these sequences is as shown in Example 1.

The region of interest containing the nucleotide of interest on each Cin1 template is amplified as in Example 1 using amplification primers 1 and 2. Primer 1 is biotinylated at the 5' end.

Formation of the probe-template hybrid is carried out as in Example 1. However, the reaction buffer contains:

1.0 µl 25 µM SF-ddTTP-526
1.0 µl 25 µM SF-ddCTP-519
1.0 µl 25 µM SF-ddGTP-505
1.0 µl 25 µM SF-ddATP-512
2.0 µl sterile water
2.0 µl 5X RT buffer
0.5 µl Reverse Transcriptase (10 µ/µl.

The labeling reaction takes place at 42°C for 10 minutes, is placed on ice and 100 µl of TE buffer is added.

The chain-terminating nucleotide which extended the probe at the position complementary to the nucleotide of interest is detected and determined as shown in Example 1. The nucleotide of interest at position 500 is identified as the nucleotide complementary to the chain-terminating nucleotide which extended the probe in the labeling reaction. The presence and nature of a polymorphism can be determined by comparing the samples tested.

From the foregoing description, one skilled in the art can easily ascertain characteristics of this invention, and without departing from the spirit and scope thereof, can make various modifications of the invention to adapt it to various uses and conditions.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: LIVAK, KENNETH J.
RAFALSKI, J. A.
SHEPHERD, NANCY S.

(ii) TITLE OF INVENTION: METHOD OF IDENTIFYING A NUCLEOTIDE PRESENT AT A DEFINED POSITION IN A NUCLEIC ACID

(iii) NUMBER OF SEQUENCES: 34

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: DU PONT COMPANY

(B) STREET: BARLEY MILL PLAZA 36/2152

(C) CITY: WILMINGTON

(D) STATE: DELAWARE

(E) COUNTRY: USA

(F) ZIP: 19880-0036

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB

(B) COMPUTER: Macintosh

(C) OPERATING SYSTEM: Macintosh System, 6.0

(D) SOFTWARE: Microsoft Word, 4.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE:

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 07/669,568

(B) FILING DATE: March 13, 1991

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: GALLEGOS, R. THOMAS

(C) REFERENCE/DOCKET NUMBER: CR-8918

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 302-892-7342

(B) TELEFAX: 302-892-7949

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTGCTGCCCG ACAACAGCAA T 21

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATTGCTGTTG TCGGGCAGCA G 21

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CAGACATTTG AGAATCAAGT GAATCG 26

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CTCGGCTCTC AGGACCATAA TCAT 24

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATGTAGAGGG CAAGCGGATC C 21

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGATCCGCTT GCCCTCTACA T 21

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CACTACTCGA TCCTCAAGCA G 21

(2) INFORMATION FOR SEQ ID NO:8:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTGCTTGAGG ATCGAGTAGT G 21

(2) INFORMATION FOR SEQ ID NO:9:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATTGCTGTTG TCGGGCAGCA G 21

(2) INFORMATION FOR SEQ ID NO:10:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGATCCGCTT GCCCTCTACA TTCTGCTGCC CGACAACAGC AAT 43

(2) INFORMATION FOR SEQ ID NO:11:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATTGCTGTTG TCGGGCAGCA GA 22

(2) INFORMATION FOR SEQ ID NO:12:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGATCCGCTT GCCCTCTACA TTCTGCTGCC CGACAACAGC AAT 43

(2) INFORMATION FOR SEQ ID NO:13:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ATTGCTGTTG TCGGGCAGCA GG 22

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGATCCGCTT GCCCTCTACA TCCTGCTGCC CGACAACAGC AAT 43

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATTGCTGTTG TCGGGCAGCA GAATGTAGAG GGCAAGCGGA TCC 43

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGATCCGCTT GCCCTCTACA T 21

(2) INFORMATION FOR SEQ ID NO:17:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ATTGCTGTTG TCGGGCAGCA GAATGTAGAG GGCAAGCGGA TCC 43

(2) INFORMATION FOR SEQ ID NO:18:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GGATCCGCTT GCCCTCTACA TT 22

(2) INFORMATION FOR SEQ ID NO:19:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ATTGCTGTTG TCGGGCAGCA GGATGTAGAG GGCAAGCGGA TCC 43

(2) INFORMATION FOR SEQ ID NO:20:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGATCCGCTT GCCCTCTACA TC 22

(2) INFORMATION FOR SEQ ID NO:21:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GGATCCGCTT GCCCTCTACA TTCTGCTGCC CGACAACAGC AAT 43

(2) INFORMATION FOR SEQ ID NO:22:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GGATCCGCTT GCCCTCTACA TTTTCTGCTG CCCGACAACA GCAAT 45

(2) INFORMATION FOR SEQ ID NO:23:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TGCTGTTGTC GGGCAGCAGA AT 22

(2) INFORMATION FOR SEQ ID NO:24:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGATCCGCTT GCCCTCTACA TTCTGCTGCC CGACAACAGC AAT 43

(2) INFORMATION FOR SEQ ID NO:25:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TGCTGTTGTC GGGCAGCAGA AA 22

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 45 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GGATCCGCTT GCCCTCTACA TTTTCTGCTG CCCGACAACA GCAAT 45

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 603 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CAGACATTTG AGAATCAAGT GAATCGTATT CTCAATGATG CTCGAGACAA AACTGGCTCC 60

TCTGCACAGA AATCCCTCTC TGAATATAAC AACTTCAAGT CTATGGTGGT GTCTGGAGCC 120

AAGGGTTCCA AGATCAACAT CTCCCAGGCA AGATGCTTCA TTTTCCAGAT ATGTGGCCTA 180

TACCAGAGTT TGTAAAGAGG ATGGTATGTA CATGTTTTGG TGTGAGGAAA GATGGAAAAA 240

ATAGTAGGGA ATTGTCACCA CCACCACCAC TGCTGCAGTG TCATGGCTTG AAACAAGATT 300

CACTCACGTG TAAAAGACCT TTTTTAAAAC AAAACAAAAC ATGGTTTTGC TGTGTAGCCC 360

AGGTTGAGTG TGAACTTTGT ATCTTTCTGC CTCCTCTTTC CAACTTTAGG TTTCAGGCAT 420

GCACTATTTC TGCCATAAAT TCATACTTTT AATGCTAGGG GAAATCATAT GCAGCCTTTC 480

CCCCCCCTTA GGTCATTGCT GTTGTCGGGC AGCAGAATGT AGAGGGCAAG CGGATCCCAT 540

TTGGATTCAA GCATCGGACT CTTCCTCACT TTATCAAGGA TGATTATGGT CCTGAGAGCC 600

GAG 603

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 603 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CTCGGCTCTC AGGACCATAA TCATCCTTGA TAAAGTGAGG AAGAGTCCGA TGCTTGAATC 60

CAAATGGGAT CCGCTTGCCC TCTACATTCT GCTGCCCGAC AACAGCAATG ACCTAAGGGG 120

GGGGAAAGGC TGCATATGAT TTCCCCTAGC ATTAAAAGTA TGAATTTATG GCAGAAATAG 180

TGCATGCCTG AAACCTAAAG TTGGAAAGAG GAGGCAGAAA GATACAAAGT TCACACTCAA 240

CCTGGGCTAC ACAGCAAAAC CATGTTTTGT TTTGTTTTAA AAAAGGTCTT TTACACGTGA 300

GTGAATCTTG TTTCAAGCCA TGACACTGCA GCAGTGGTGG TGGTGGTGAC AATTCCCTAC 360

TATTTTTTCC ATCTTTCCTC ACACCAAAAC ATGTACATAC CATCCTCTTT ACAAACTCTG 420

GTATAGGCCA CATATCTGGA AAATGAAGCA TCTTGCCTGG GAGATGTTGA TCTTGGAACC 480

CTTGGCTCCA GACACCACCA TAGACTTGAA GTTGTTATAT TCAGAGAGGG ATTTCTGTGC 540

AGAGGAGCCA GTTTTGTCTC GAGCATCATT GAGAATACGA TTCACTTGAT TCTCAAATGT 600

CTG 603

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CCCCCCCTTA GGTCATTGCT GTTGTCGGGC AGCAGAATGT AGAGGGCAAG CGGATCCCAT 60

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

ATGGGATCCG CTTGCCCTCT ACATTCTGCT GCCCGACAAC AGCAATGACC TAAGGGGGGG 60

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CCCCCCCTTA GGTCATTGCT GTTGTCGGGC AGCAGGATGT AGAGGGCAAG CGGATCCCAT 60

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ATGGGATCCG CTTGCCCTCT ACATCCTGCT GCCCGACAAC AGCAATGACC TAAGGGGGGG 60

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 295 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GTCGTGCGAG GAGCAGACGT AGCGCCCGGC CGCGGCCGGG TTCTCGAAGA GGAAGATCTC 60

GGCGTCGCAG AGGTCGTCGA GGTGGATGAG CTGCACCTGC TTGAGGATCG AGTAGTGCGG 120

CGCGTTCCCC GTGATGAGCG CCAGCGCGGT GATGAGGCTG GGCGGCATGG ACGCGCTGAT 180

GAACGGGCCG ACCACGAGCG TCGGGATGAT GGTGACCAGG TCCAGGCCGT GCTCCGCCGC 240

GTACGCCAGG GCCGCCTTCT CCGCCAGGGT TTTAGACACG AAGTACATCT GCAGG 295

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 295 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CCTGCAGATG TACTTCGTGT CTAAAACCCT GGCGGAGAAG GCGGCCCTGG CGTACGCGGC 60

GGAGCACGGC CTGGACCTGG TCACCATCAT CCCGACGCTC GTGGTCGGCC CGTTCATCAG 120

CGCGTCCATG CCGCCCAGCC TCATCACCGC GCTGGCGCTC ATCACGGGGA ACGCGCCGCA 180

CTACTCGATC CTCAAGCAGG TGCAGCTCAT CCACCTCGAC GACCTCTGCG ACGCCGAGAT 240

CTTCCTCTTC GAGAACCCGG CCGCGGCCGG GCGCTACGTC TGCTCCTCGC ACGAC 295

What is claimed is:

1. A method of identifying a nucleotide of interest present at a defined position in a nucleic acid analyte, comprising:

a) contacting the nucleic acid analyte with a probe such that annealing takes place adjacent to the nucleotide of interest to form a hybrid;

b) contacting the hybrid with four chain terminating nucleotides;

c) extending the probe in the direction of the nucleotide of interest by addition of the chain terminating nucleotide complementary to the nucleotide of interest;

d) determining which chain terminating nucleotide was added; and

e) identifying the nucleotide of interest as the nucleotide complementary to the chain terminating nucleotide which was added.

2. The method of Claim 1 wherein the nucleic acid analyte is a sequence of DNA or RNA.

3. The method of Claim 1 wherein the probe is a sequence of DNA or RNA.

4. The method of Claim 1 wherein the chain terminating nucleotides are dideoxynucleotides.

5. The method of Claim 1 wherein the probe is extended chemically.

6. The method of Claim 1 wherein the probe is extended enzymatically.

7. The method of Claim 1 wherein the added chain terminating nucleotide is determined by detecting the presence of a signal generator.

8. The method of Claim 1 wherein the nucleic acid analyte is single stranded.

9. The method of Claim 1 wherein the nucleic acid analyte is immobilized on a solid support.

10. The method of Claim 1 wherein the probe is immobilized on a solid support.

11. A kit for identification of a nucleotide of interest in a nucleic acid analyte, comprising:

a) a probe which comprises a primer sequence complementary to the nucleic acid analyte and capable of binding the nucleic acid analyte with sufficient specificity to form a stable hybrid adjacent to the nucleotide of interest;

b) a plurality of reporter labeled chain terminating nucleotide triphosphates; and

c) a primer-dependent nucleic acid polymerase.

12. A method of identifying a nucleotide of interest present at a defined position in a nucleic acid analyte, comprising:

a) contacting the nucleic acid analyte with a probe such that annealing takes place adjacent to the nucleotide of interest;

b) contacting the hybrid with at least one chain terminating nucleotide;

c) extending the probe in the direction of the nucleotide of interest by addition of the chain terminating nucleotide complementary to the nucleotide of interest;

d) determining which chain terminating nucleotide was added; and

e) identifying the nucleotide of interest as the nucleotide complementary to the chain terminating nucleotide which was added.

13. The method of Claim 12 wherein the nucleic acid analyte is a sequence of DNA or RNA.

14. The method of Claim 12 wherein the probe is a sequence of DNA or RNA.

15. The method of Claim 12 wherein the chain terminating nucleotides are dideoxynucleotides.

16. The method of Claim 12 wherein the probe is extended chemically.

17. The method of Claim 12 wherein the probe is extended enzymatically.

18. The method of Claim 12 wherein the added chain terminating nucleotide is determined by detecting the presence of a signal generator.

19. The method of Claim 12 wherein the nucleic acid analyte is single stranded.

20. The method of Claim 12 wherein the nucleic acid analyte is immobilized on a solid support.

21. The method of Claim 12 wherein the probe is immobilized on a solid support.

22. A kit for identification of a nucleotide of interest in a nucleic acid analyte, comprising:

a) a probe which comprises a primer sequence complementary to the nucleic acid analyte and capable of binding the nucleic acid analyte with sufficient specificity to form a stable hybrid adjacent to the nucleotide of interest;

b) at least one of reporter labeled chain terminating nucleotide triphosphates; and

c) a primer-dependent nucleic acid polymerase.

FIG. 1A

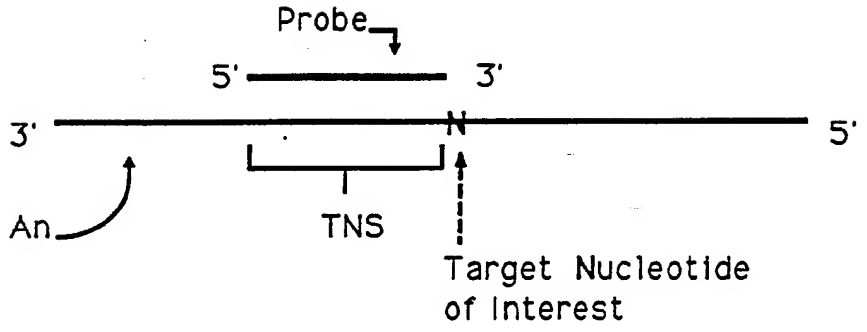

Probe
5'
3'
3'
N
5'
An
TNS
Target Nucleotide
of Interest

FIG. 1B

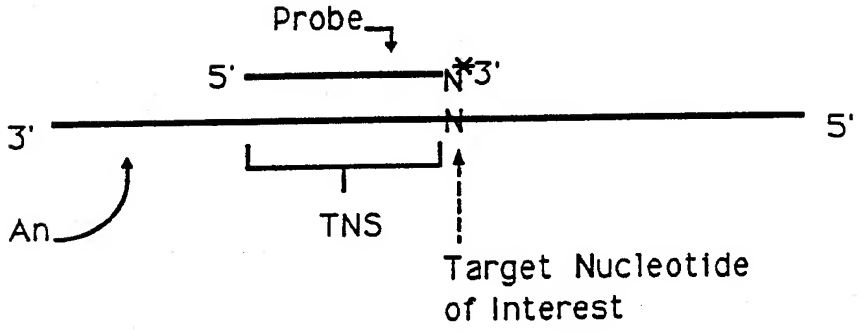

Probe
5'
N*3'
3'
N
5'
An
TNS
Target Nucleotide
of Interest

FIG. 1C

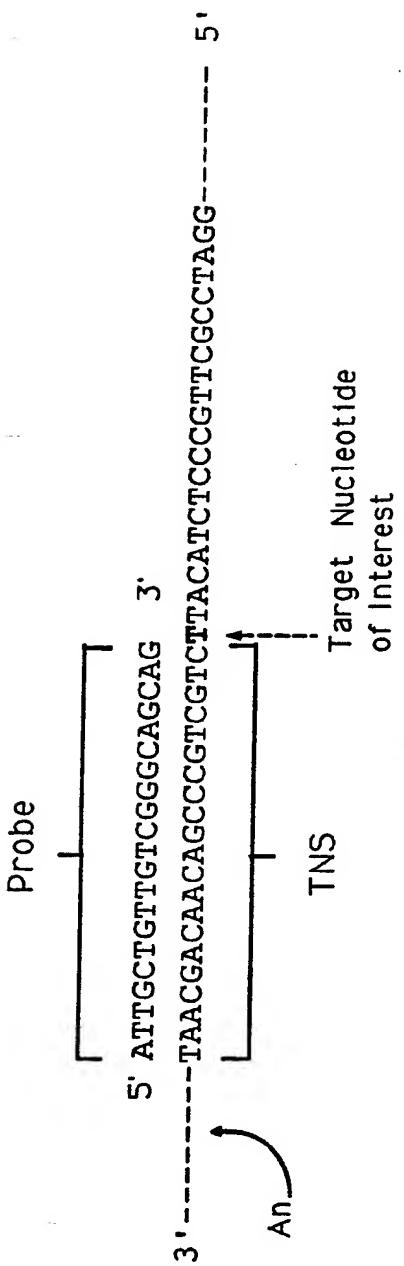

Probe
5' ATTGCTGTTGTCGGGCAGCAG 3'
3'----------TAACGACAACAGCCCGTCGTCTTACATCTCCCGTTCGCCTAGG------- 5'
An
TNS
Target Nucleotide
of Interest

FIG. 1D

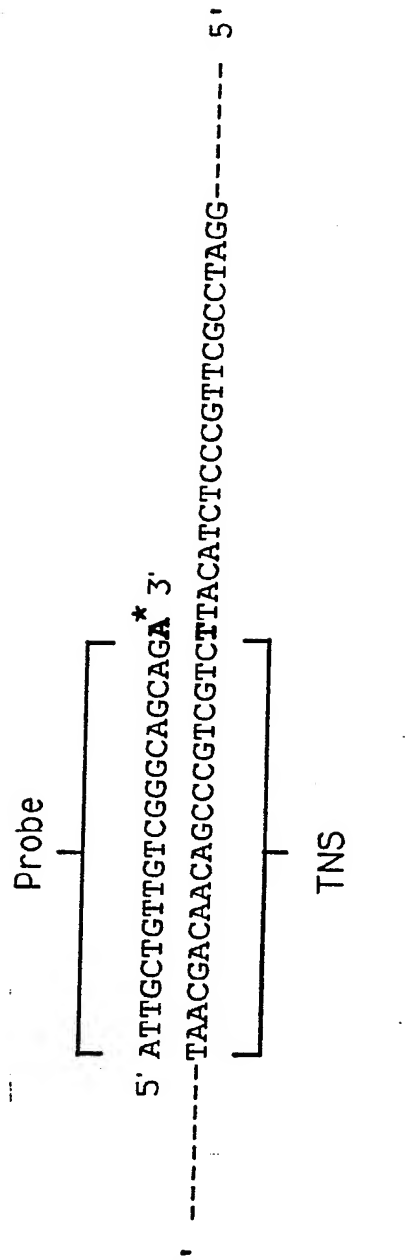

Probe
5' ATTGCTGTTGTCGGGCAGCAGA* 3'
3' ----------TAACGACAACAGCCCGTCGTCTTACATCTCCCGTTCGCCTAGG------- 5'
TNS

FIG. 1E

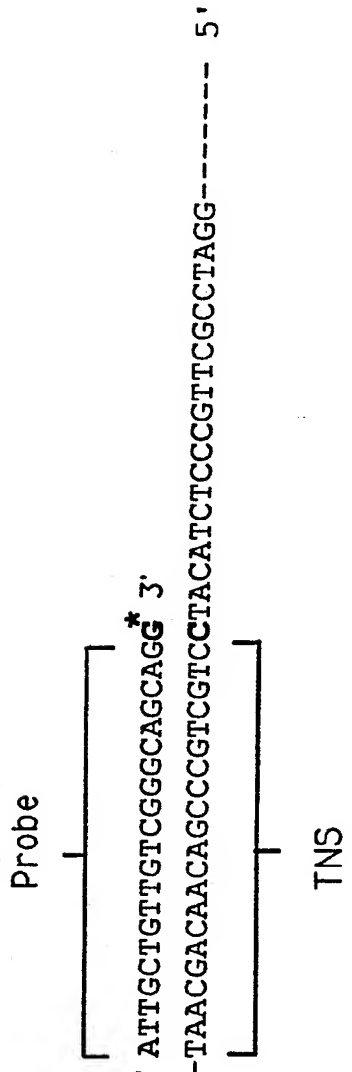

Probe
TNS
5' ATTGCTGTTGTCGGGCAGCAGG* 3'
3'--------TAACGACAACAGCCCGTCGTCCTACATCTCCCGTTCGCCTAGG------- 5'

FIG. 1F

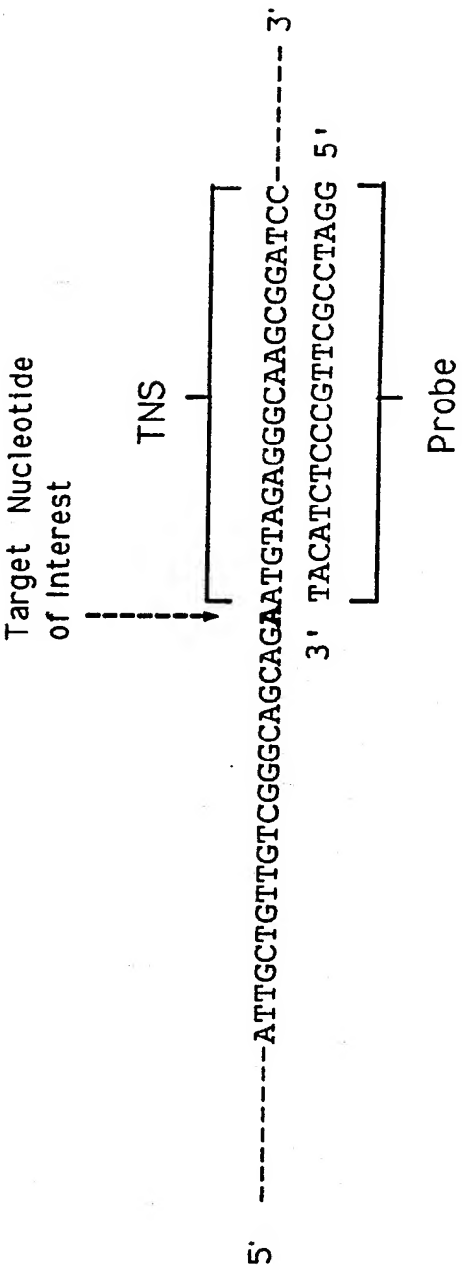

Target Nucleotide of Interest
TNS
Probe
5' --------ATTGCTGTTGTCGGGCAGCAGAATGTAGAGGGCAAGCGGATCC------- 3'
3' TACATCTCCCGTTCGCCTAGG 5'

FIG.1G

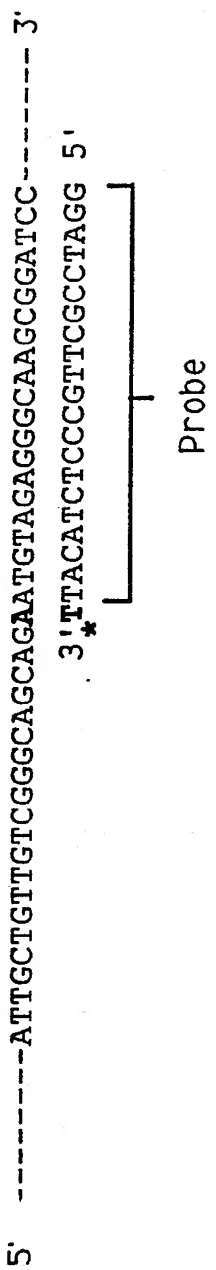

TNS
5' ---------ATTGCTGTTGTCGGGCAGCAGAATGTAGAGGGCAAGCGGATCC------ 3'
3'*TTACATCTCCCGTTCGCCTAGG 5'
Probe

FIG.1H

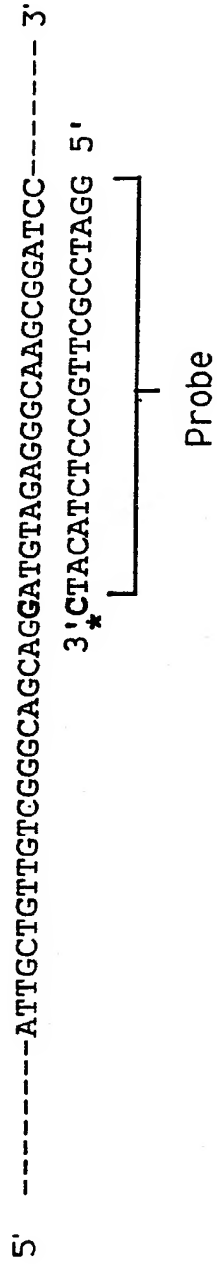

TNS
5' ---------ATTGCTGTTGTCGGGCAGCAGGATGTAGAGGGCAAGCGGATCC------ 3'
3'*CTACATCTCCCGTTCGCCTAGG 5'
Probe

FIG.2

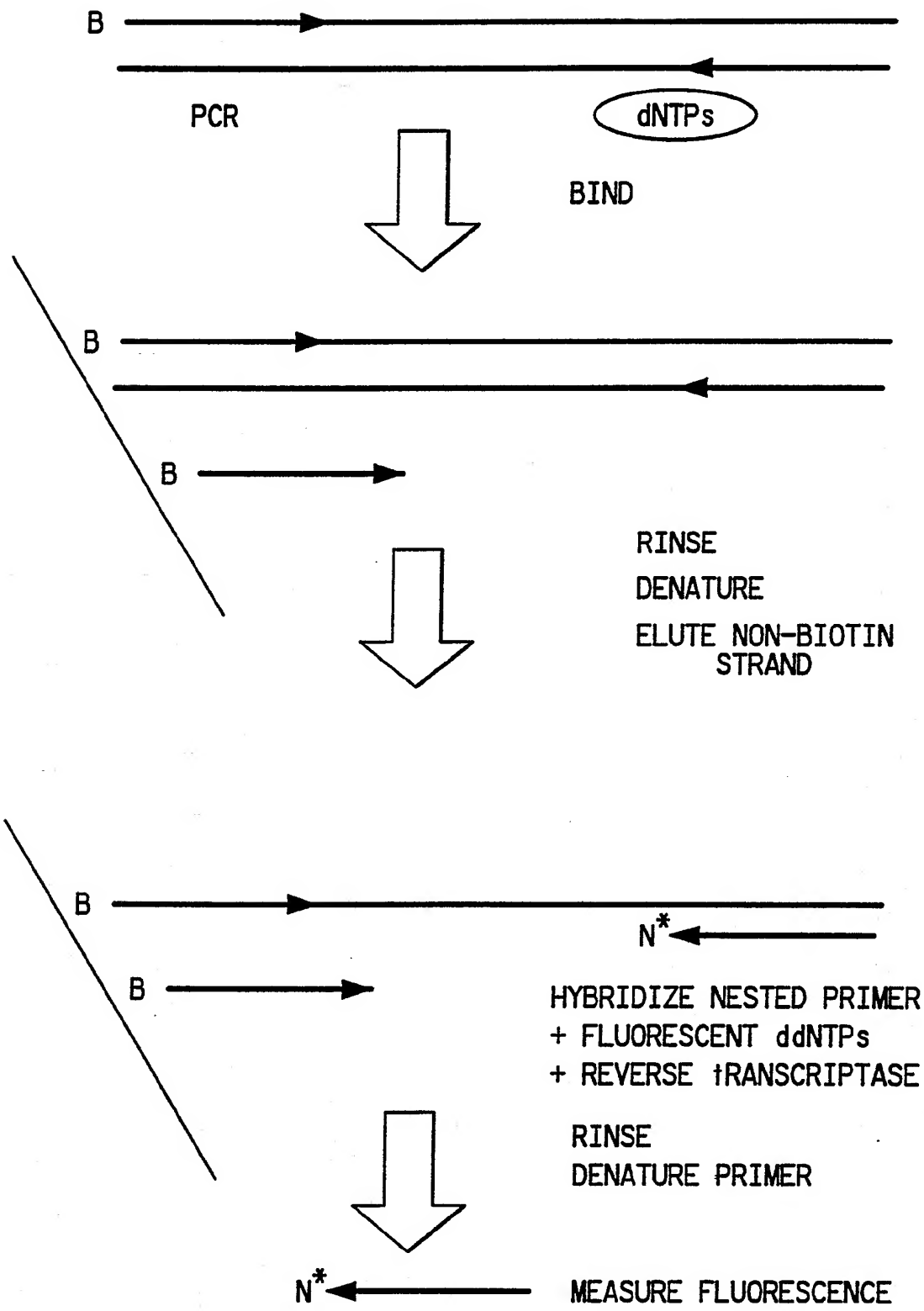

B
PCR
dNTPs
BIND
B
B
RINSE
DENATURE
ELUTE NON-BIOTIN STRAND
B
N*
B
HYBRIDIZE NESTED PRIMER
+ FLUORESCENT ddNTPs
+ REVERSE tRANSCRIPTASE
RINSE
DENATURE PRIMER
N*
MEASURE FLUORESCENCE

FIG.3

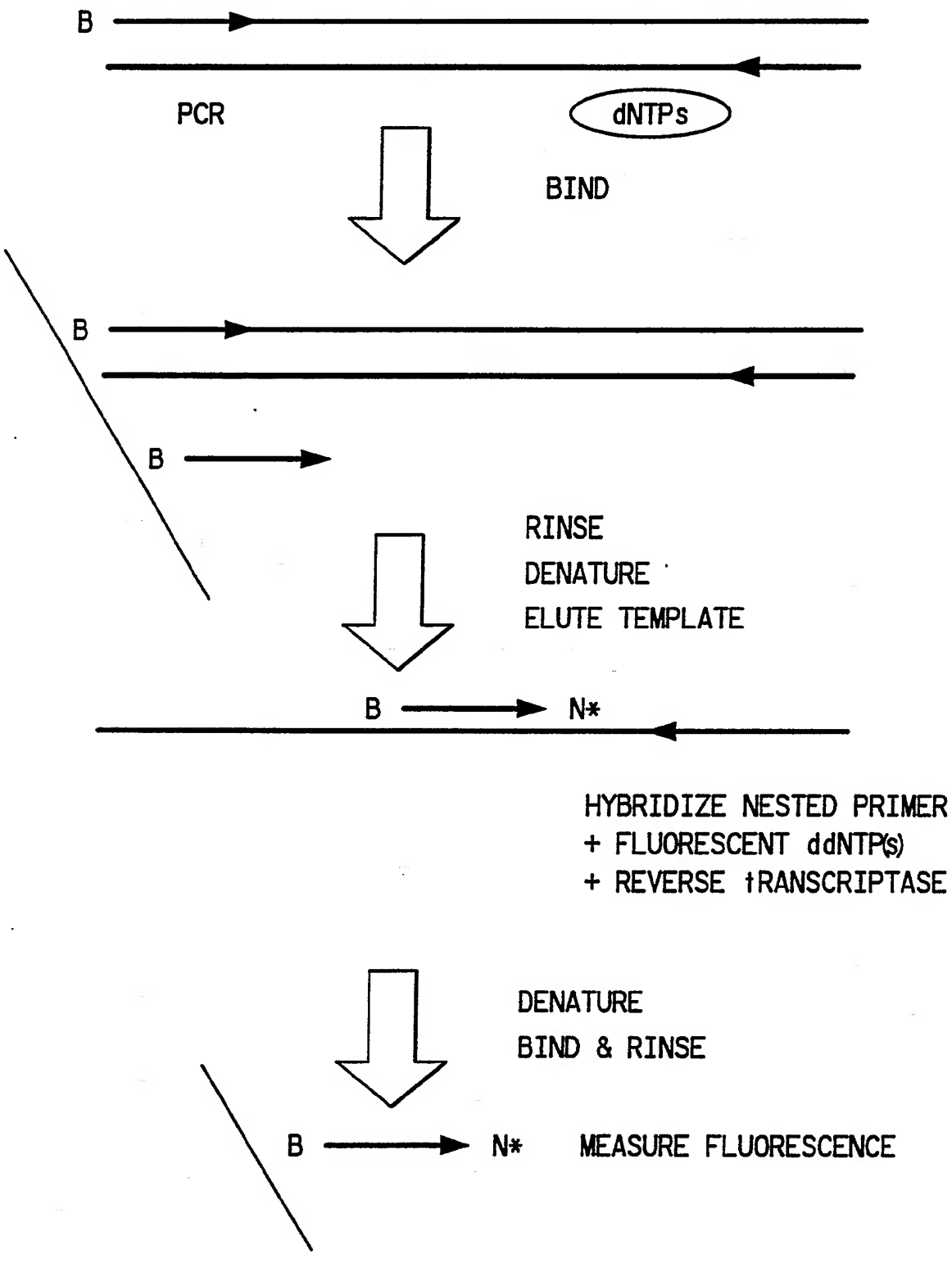

B
PCR
dNTPs
BIND
B
B
RINSE
DENATURE
ELUTE TEMPLATE
B
N*
HYBRIDIZE NESTED PRIMER
+ FLUORESCENT ddNTP(s)
+ REVERSE tRANSCRIPTASE
DENATURE
BIND & RINSE
B
N*
MEASURE FLUORESCENCE

FIG.4

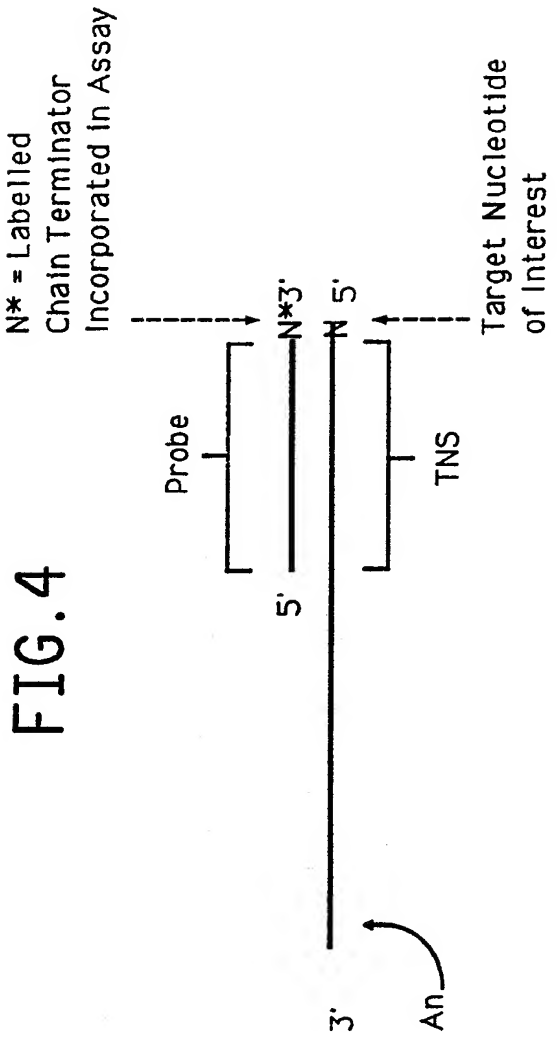


FIG.5A

3' --------TAACGACAACAGCCCGTCGTCTTACATCTCCCGTTCGCCTAGG------- 5'

versus

3' --------TAACGACAACAGCCCGTCGTCTTTTACATCTCCCGTTCGCCTAGG----- 5'

FIG.5B

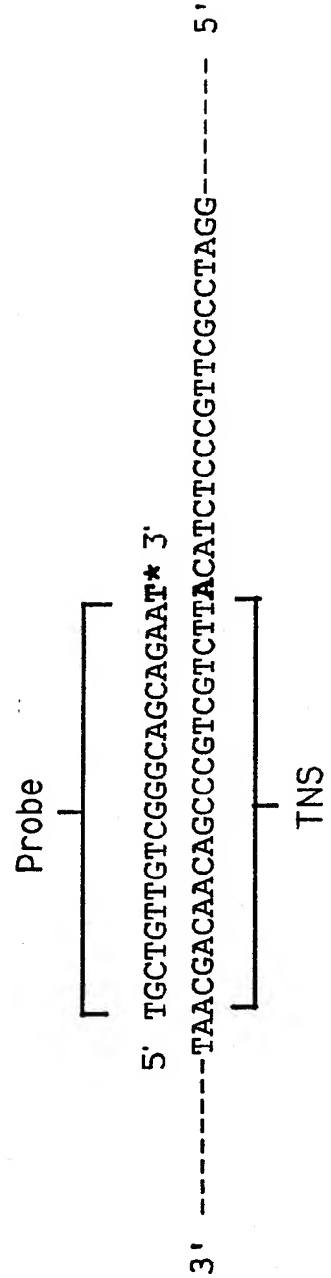

Probe
TNS
5' TGCTGTTGTCGGGCAGCAGAAT* 3'
3' ---------TAACGACACAGCCCGTCGTCTTACATCTCCCGTTCGCCTAGG------- 5'

FIG.5C

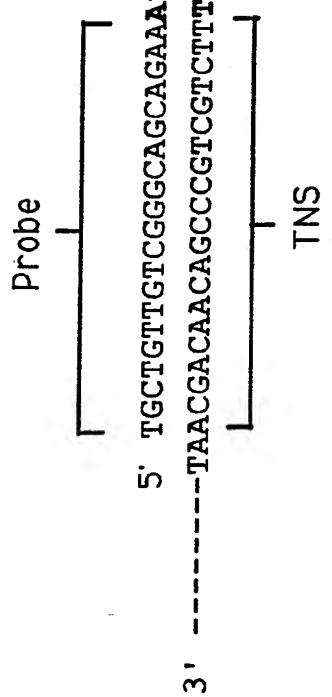

Probe
TNS
5' TGCTGTTGTCGGGCAGCAGAAA* 3'
3' ---------TAACGACACAGCCCGTCGTCTTTACATCTCCCGTTCGCCTAGG----- 5'

FIG.6

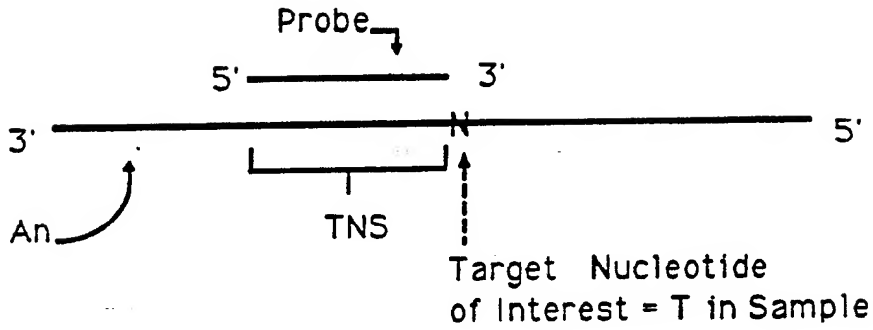


FIG. 7A

GREEN TO RED RATIO FOR FLUORESCENT NUCLEOTIDES

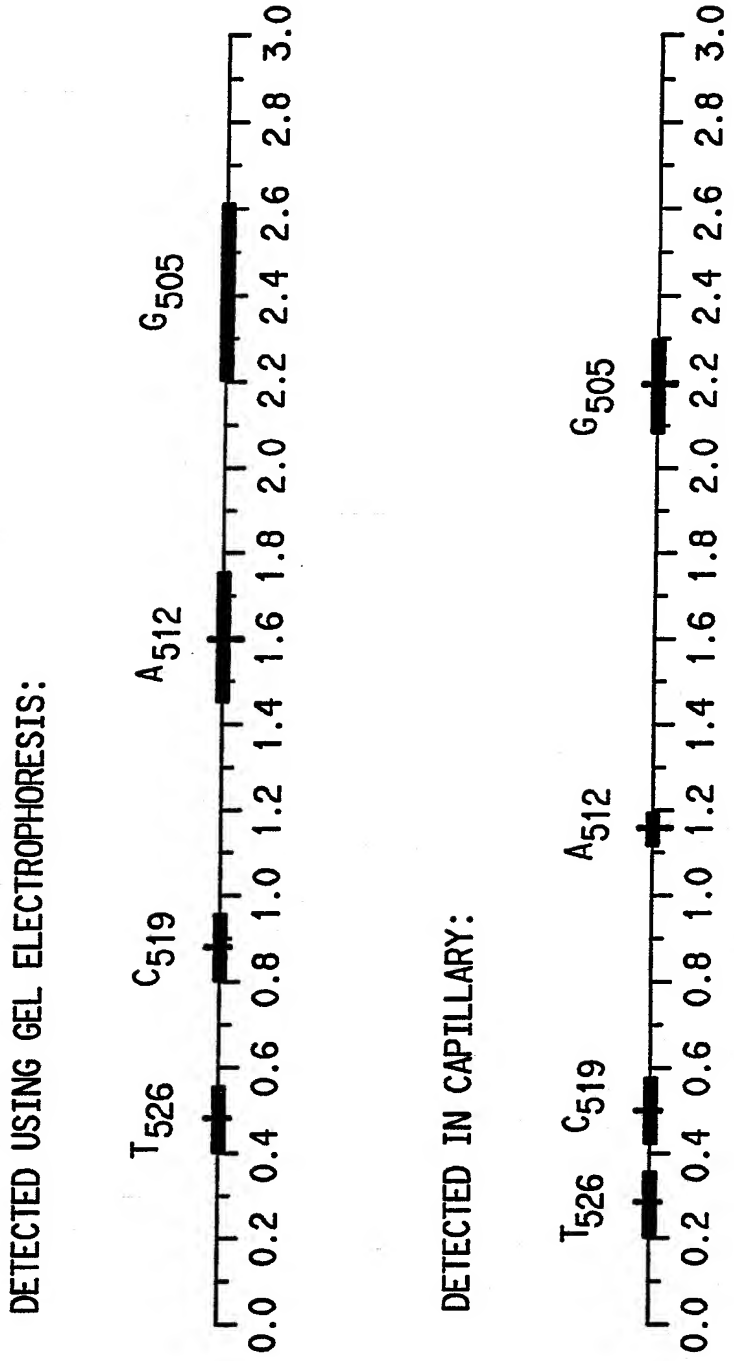


FIG.7B

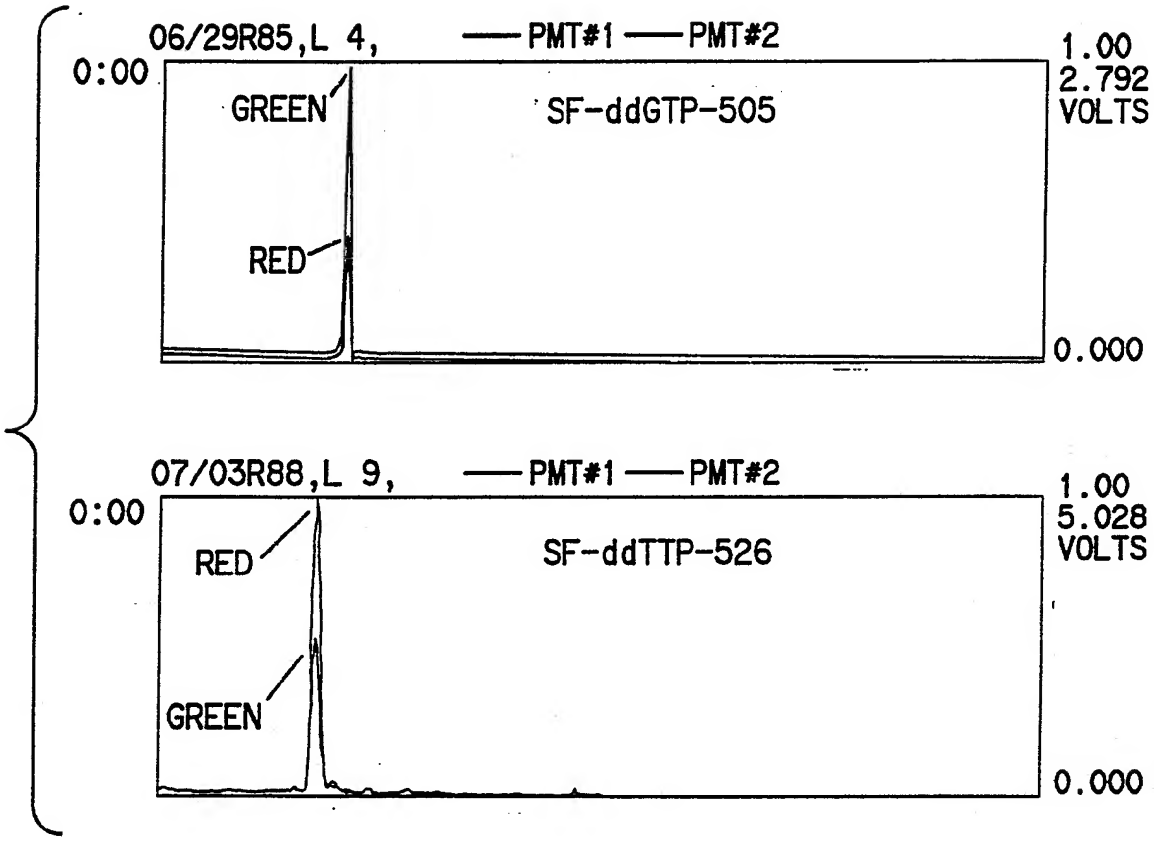

06/29R85,L 4,
PMT#1
PMT#2
0:00
GREEN
RED
SF-ddGTP-505
1.00
2.792
VOLTS
0.000
07/03R88,L 9,
PMT#1
PMT#2
0:00
RED
GREEN
SF-ddTTP-526
1.00
5.028
VOLTS
0.000

FIG.7C

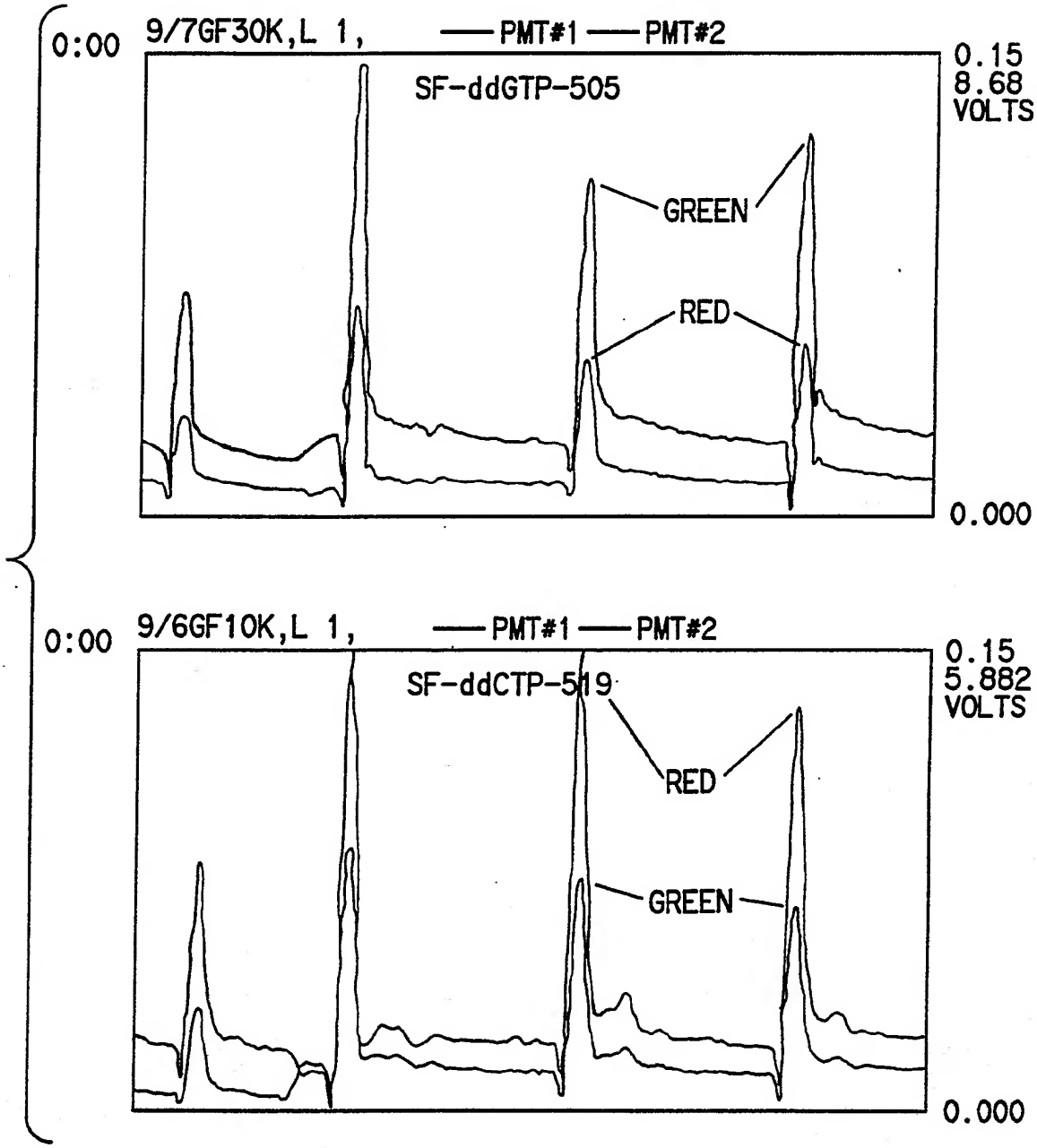

0:00
9/7GF30K,L 1,
PMT#1
PMT#2
SF-ddGTP-505
GREEN
RED
0.15
8.68
VOLTS
0.000
0:00
9/6GF10K,L 1,
PMT#1
PMT#2
SF-ddCTP-519
RED
GREEN
0.15
5.882
VOLTS
0.000

FIG.8A

PCR PRIMER 1

CAGACATTTGAGAATCAAGTGAATCGTATTCTCAATGATGCTCGAGACAAAACTGGCTCC

```
4915 -----+---------+---------+---------+---------+---------+---------+---- 4974
```

GTCTGTAAACTCTTAGTTCACTTAGCATAAGAGTTACTACGAGCTCTGTTTTGACCGAGG

TCTGCACAGAAATCCCTCTCTGAATATAACAACTTCAAGTCTATGGTGGTGTCTGGAGCC

```
4975 -----+---------+---------+---------+---------+---------+---------+---- 5034
```

AGACGTGTCTTTAGGGAGAGACTTATATTGTTGAAGTTCAGATACCACCACAGACCTCGG

AAGGGTTCCAAGATCAACATCTCCCAGGCAAGATGCTTCATTTTCCAGATATGTGGCCTA

```
5035 -----+---------+---------+---------+---------+---------+---------+---- 5094
```

TTCCCAAGGTTCTAGTTGTAGAGGGTCCGTTCTACGAAGTAAAAGGTCTATACACCGGAT

TACCAGAGTTTGTAAAGAGGATGGTATGTACATGTTTTGGTGTGAGGAAAGATGGAAAAA

```
5095 -----+---------+---------+---------+---------+---------+---------+---- 5154
```

ATGGTCTCAAACATTTCTCCTACCATACATGTACAAAACCACACTCCTTTCTACCTTTTT

ATAGTAGGGAATTGTCACCACCACCACCACTGCTGCAGTGTCATGGCTTGAAACAAGATT

```
5155 -----+---------+---------+---------+---------+---------+---------+---- 5214
```

TATCATCCCTTAACAGTGGTGGTGGTGGTGACGACGTCACAGTACCGAACTTTGTTCTAA

CACTCACGTGTAAAAGACCTTTTTTAAAACAAAACAAAACATGGTTTTGCTGTGTAGCCC

```
5215 -----+---------+---------+---------+---------+---------+---------+---- 5274
```

GTGAGTGCACATTTTCTGGAAAAAATTTTGTTTTGTTTTGTACCAAAACGACACATCGGG

AGGTTGAGTGTGAACTTTGTATCTTTCTGCCTCCTCTTTCCAACTTTAGGTTTCAGGCAT

```
5275 -----+---------+---------+---------+---------+---------+---------+---- 5334
```

TCCAACTCACACTTGAAACATAGAAAGACGGAGGAGAAAGGTTGAAATCCAAAGTCCGTA

CONTINUED ON FIG. 8B

FIG.8B

CONTINUED FROM FIG. 8A

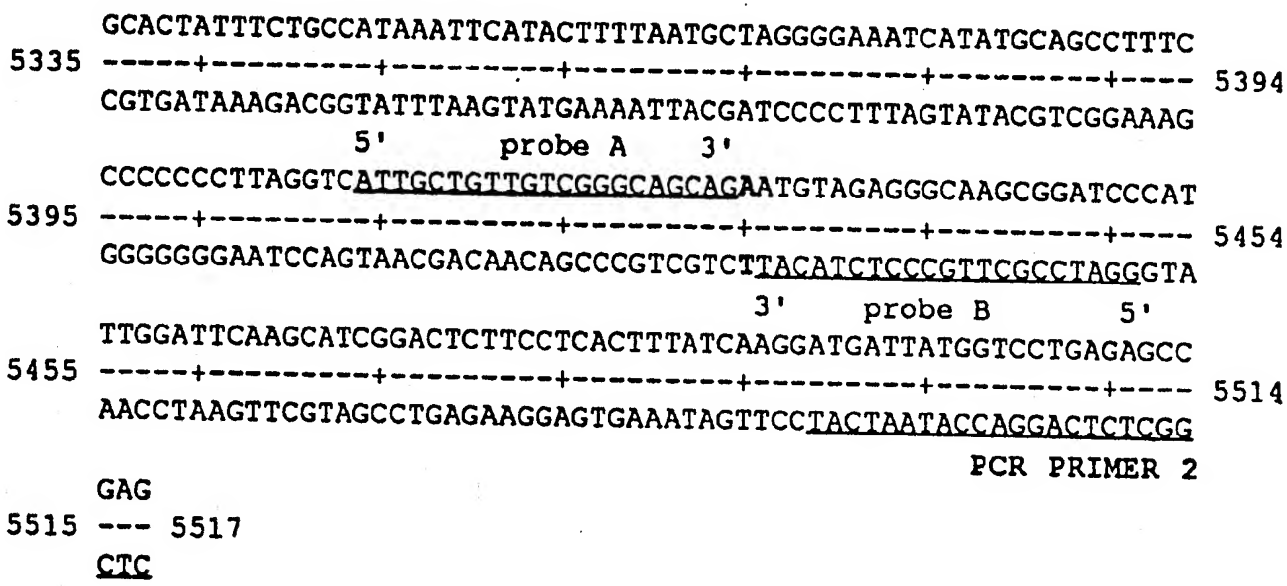
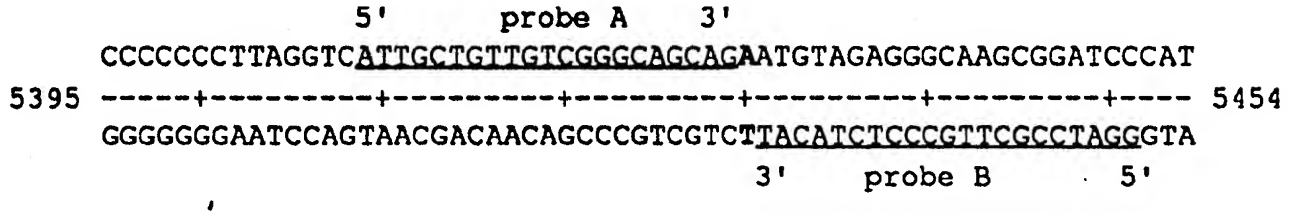


FIG.9

WILDTYPE ALLELE



MUTANT ALLELE

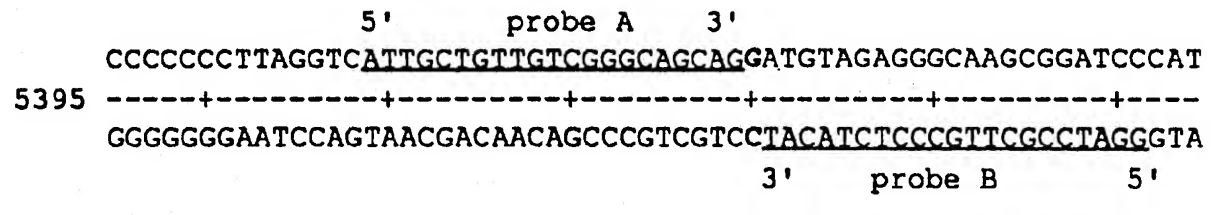


FIG.10A

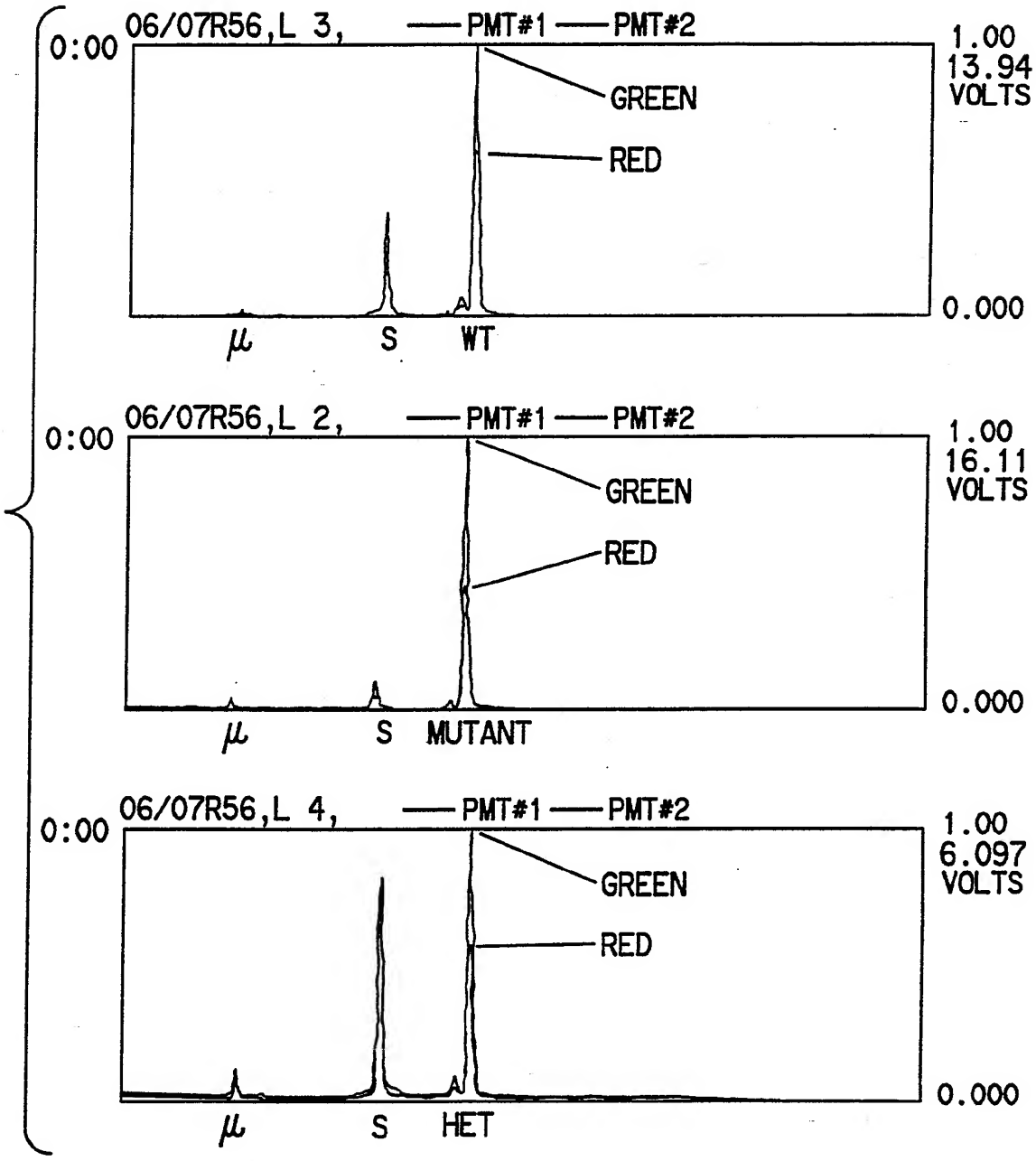

0:00
06/07R56,L 3,
PMT#1
PMT#2
1.00
13.94
VOLTS
GREEN
RED
0.000
μ
S
WT
0:00
06/07R56,L 2,
PMT#1
PMT#2
1.00
16.11
VOLTS
GREEN
RED
0.000
μ
S
MUTANT
0:00
06/07R56,L 4,
PMT#1
PMT#2
1.00
6.097
VOLTS
GREEN
RED
0.000
μ
S
HET

FIG. 10B

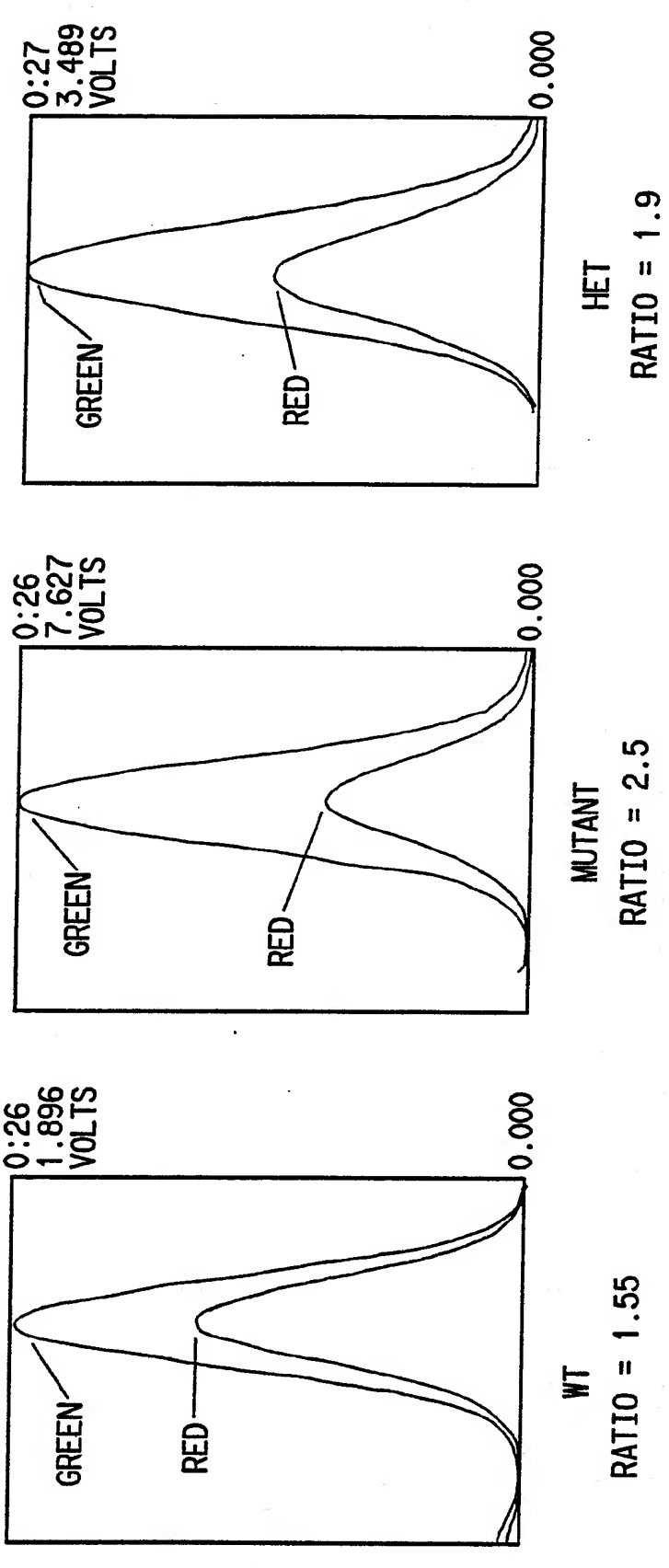

0:27
3.489
VOLTS
0.000
GREEN
RED
HET
RATIO = 1.9
0:26
7.627
VOLTS
0.000
GREEN
RED
MUTANT
RATIO = 2.5
0:26
1.896
VOLTS
0.000
GREEN
RED
WT
RATIO = 1.55

FIG.11

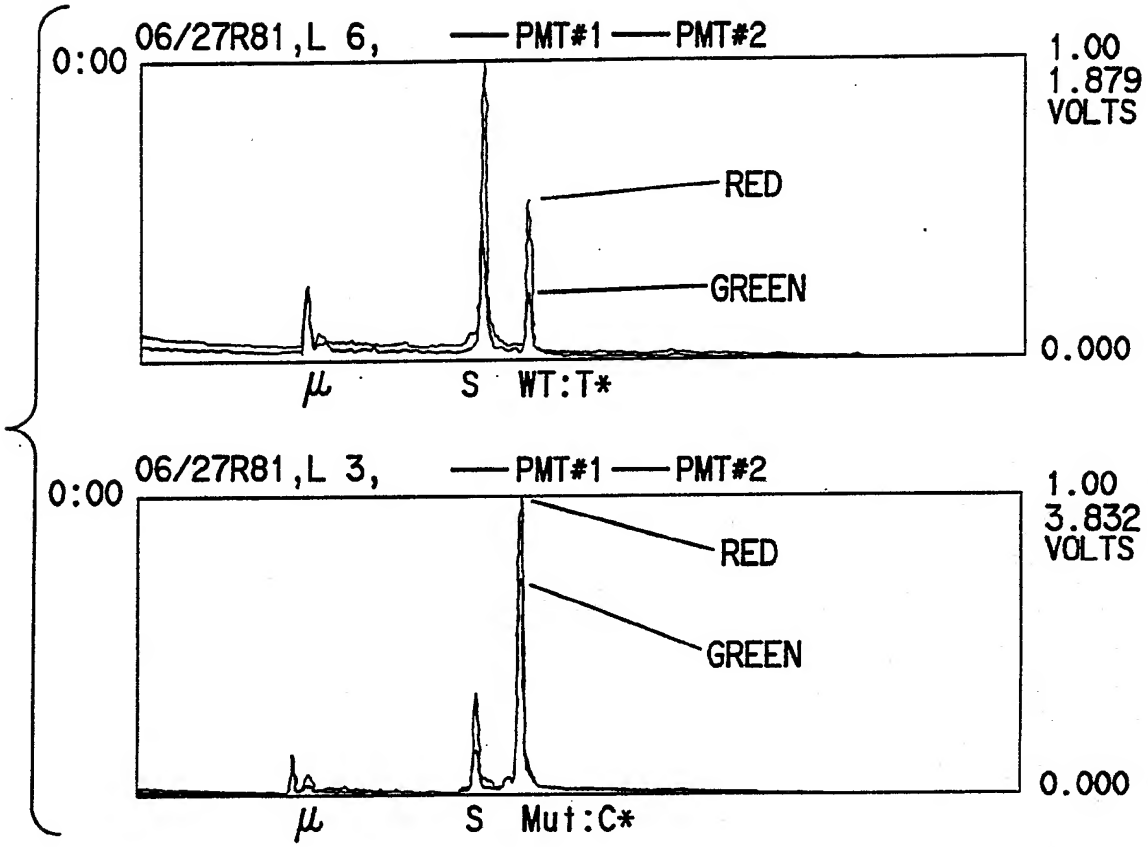

0:00
06/27R81,L 6,
PMT#1
PMT#2
1.00
1.879
VOLTS
RED
GREEN
0.000
μ
S
WT:T*
0:00
06/27R81,L 3,
PMT#1
PMT#2
1.00
3.832
VOLTS
RED
GREEN
0.000
μ
S
Mut:C*

FIG. 12

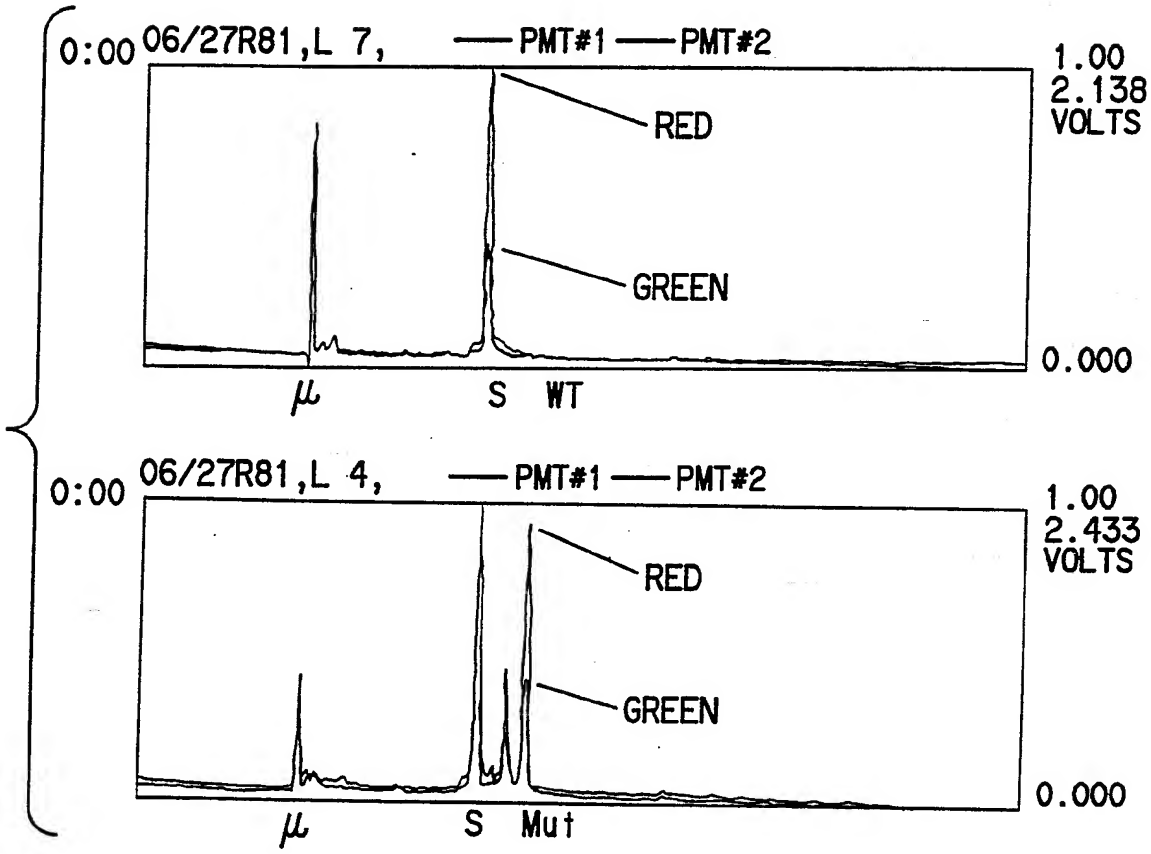

0:00
06/27R81,L 7,
PMT#1
PMT#2
1.00
2.138
VOLTS
RED
GREEN
0.000
μ
S
WT
0:00
06/27R81,L 4,
PMT#1
PMT#2
1.00
2.433
VOLTS
RED
GREEN
0.000
μ
S
Mut

FIG. 13

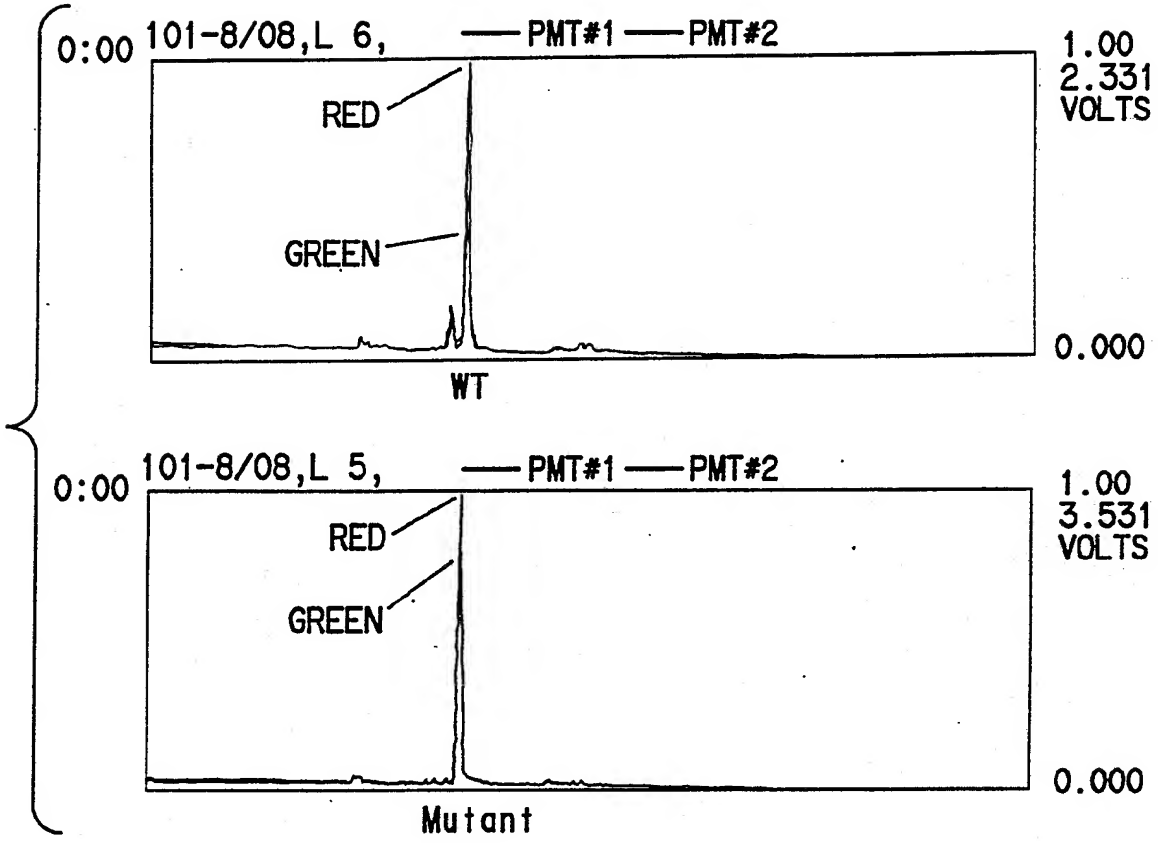

0:00
101-8/08,L 6,
PMT#1
PMT#2
1.00
2.331
VOLTS
RED
GREEN
0.000
WT
0:00
101-8/08,L 5,
PMT#1
PMT#2
1.00
3.531
VOLTS
RED
GREEN
0.000
Mutant

FIG.14

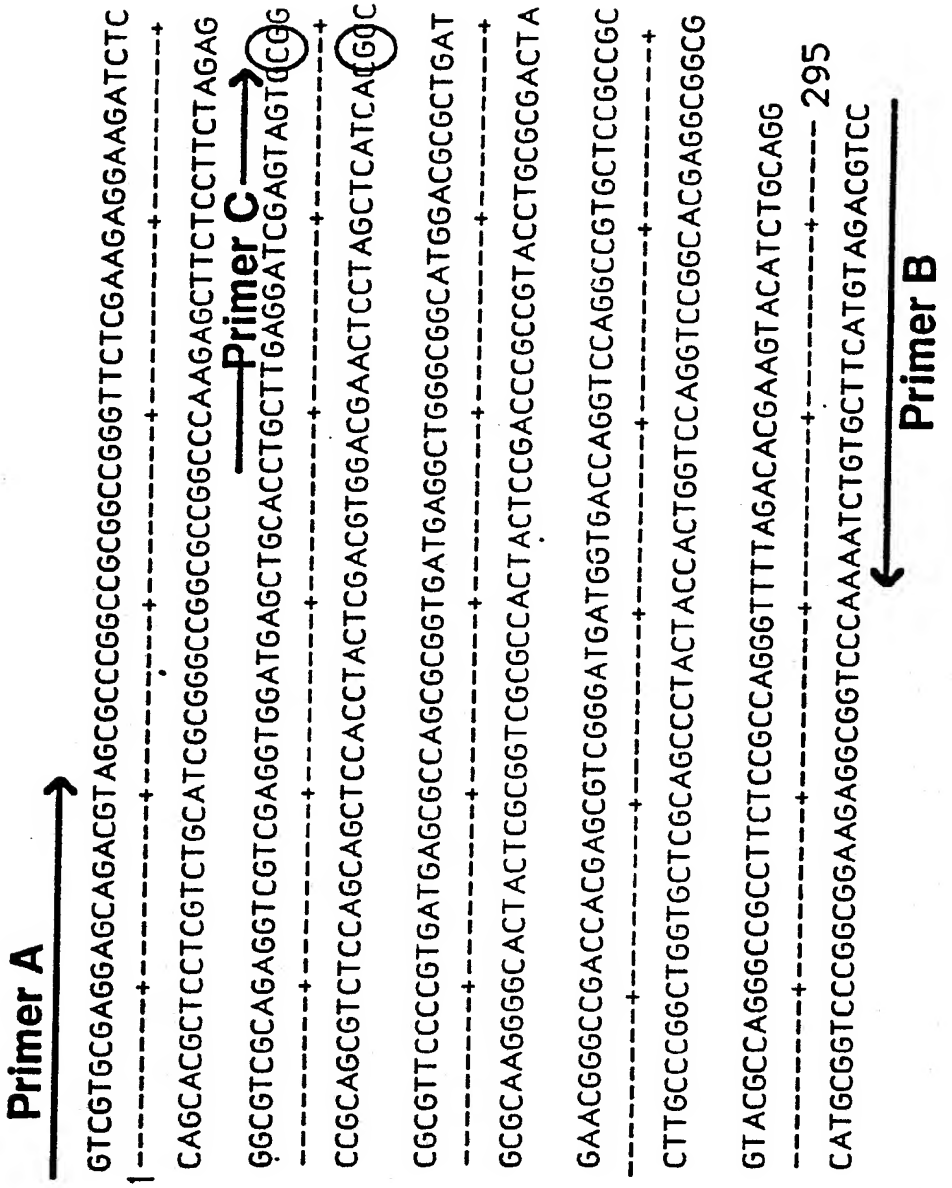

Primer A
GTCGTGCGAGGAGCAGACGTAGCGCCCGGCCGCGGCCGGGTTCTCGAAGAGGAAGATCTC
1
CAGCACGCTCCTCGTCTGCATCGCGGGCCGGCGCCGGCCCAAGAGCTTCTCCTTCTAGAG
Primer C
GGCGTCGCAGAGGTCGTCGAGGTGGATGAGCTGCACCTGCTTGAGGATCGAGTAGTCGGG
CCGCAGCGTCTCCAGCAGCTCCACCTACTCGACGTGGACGAACTCCTAGCTCATCAGCC
CGCGTTCCCCGTGATGAGCGCCAGCGCGGTGATGAGGCTGGGCGGCATGGACGCGCTGAT
GCGCAAGGGGCACTACTCGCGGTCGCGCCACTACTCCGACCCGCCGTACCTGCGCGACTA
GAACGGGCCGACCACGAGCGTCGGGATGATGGTGACCAGGTCCAGGCCGTGCTCCGCCGC
CTTGCCCGGCTGGTGCTCGCAGCCCTACTACCACTGGTCCAGGTCCGGCACGAGGCGGCG
GTACGCCAGGGCCGCCTTCTCCGCCAGGGTTTTAGACACGAAGTACATCTGCAGG
CATGCGGTCCCGGCGGAAGAGGCGGTCCCAAAATCTGTGCTTCATGTAGACGTCC
295
Primer B

FIG. 15

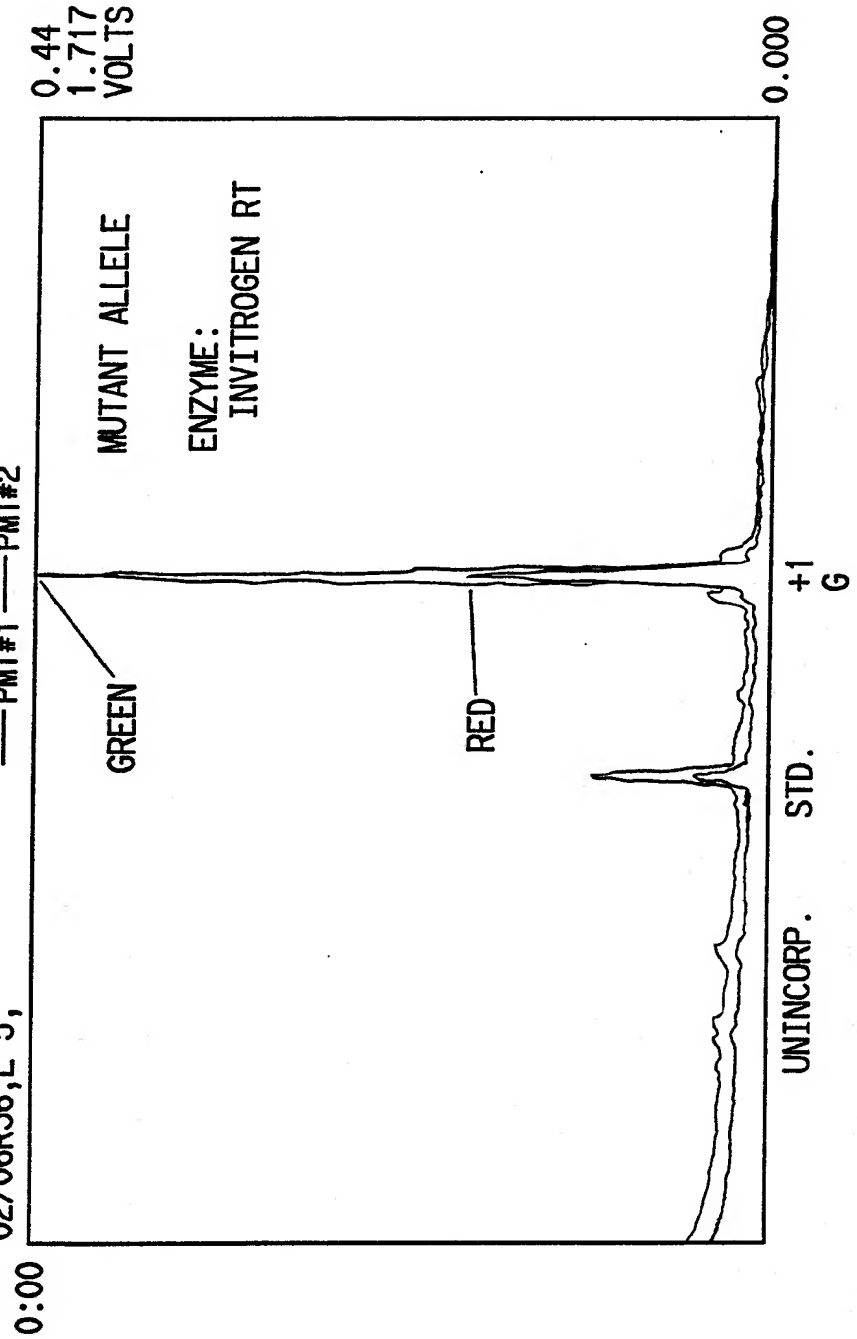
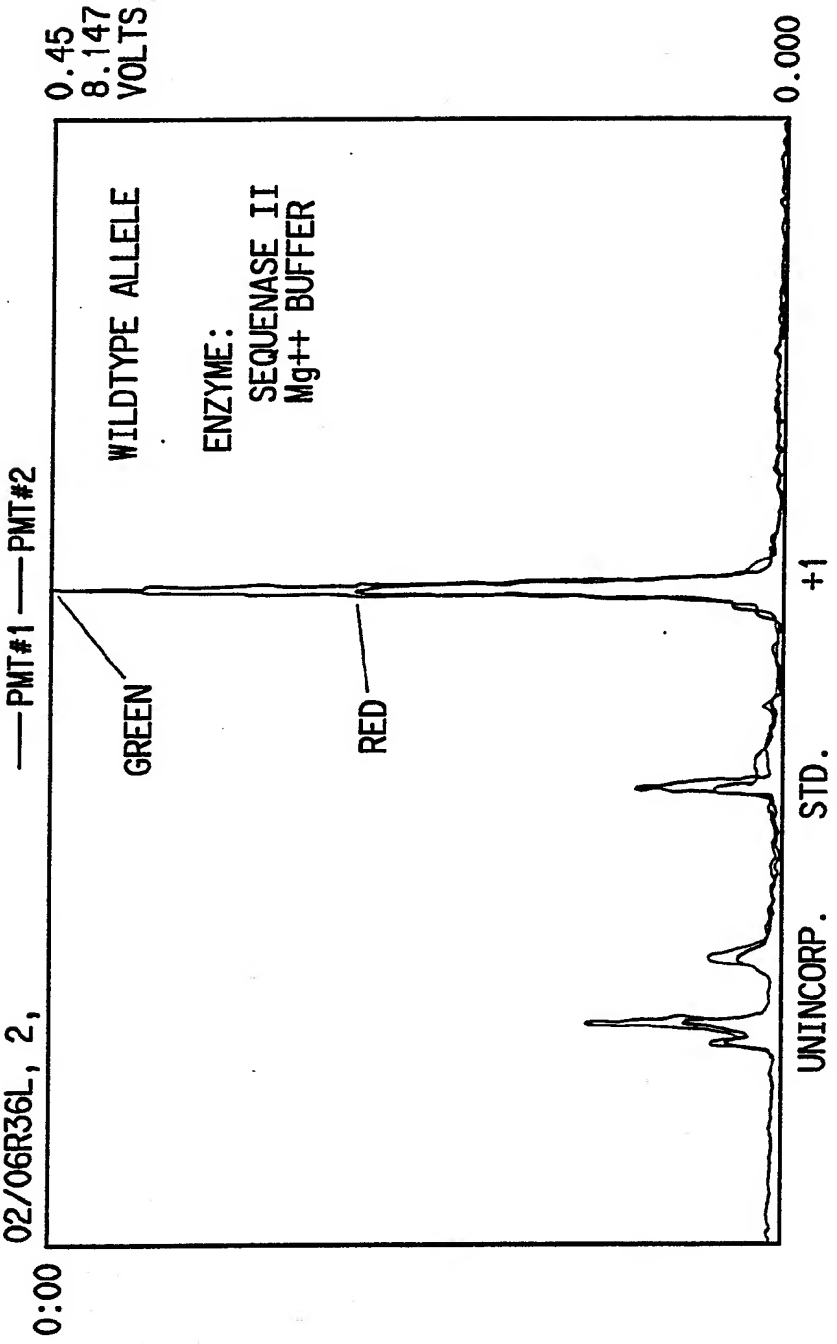

0.44
1.717
VOLTS
0.000
MUTANT ALLELE
ENZYME:
INVITROGEN RT
PMT#1
PMT#2
GREEN
RED
02/06R36,L 5,
0:00
UNINCORP.
STD.
+1
G

FIG.16


0.45
8.147
VOLTS
0.000
PMT#1
PMT#2
WILDTYPE ALLELE
ENZYME:
SEQUENASE II
Mg++ BUFFER
GREEN
RED
0:00
02/06R36L, 2,
UNINCORP.
STD.
+1

INTERNATIONAL SEARCH REPORT

International Application No PCT/US 92/01691

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.Cl. 5 C12Q1/68

II. FIELDS SEARCHED

Minimum Documentation Searched[7]

Classification System	Classification Symbols
Int.Cl. 5	C12Q

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched[8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

Category °	Citation of Document,[11] with indication, where appropriate, of the relevant passages[12]	Relevant to Claim No.[13]
X	EP,A,0 412 883 (BERTIN & CIE) 13 February 1991 see the whole document	1-22
Y	NUCLEIC ACIDS RESEARCH. vol. 18, no. 12, 1990, ARLINGTON, VIRGINIA US page 3671; B. P. SOKOLOV: 'Primer extension technique for the detection of single nucleotide in genomic DNA' cited in the application see the whole document	1-22

-/--

° Special categories of cited documents :[10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
18 JUNE 1992	30. 06. 92
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	MOLINA GALAN E.

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category °	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
Y	SCIENCE. vol. 238, 16 October 1987, LANCASTER, PA US pages 336 - 341; J. M. PROBER ET AL.: 'A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides' cited in the application see abstract	1-22
A	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA. vol. 86, April 1989, WASHINGTON US D. Y. WU ET AL.: 'Allele-specific enzymatic amplification of beta-globin genomic DNA for diagnosis of sickle cell anemia' cited in the application	
A	WO,A,8 909 283 (E. HYMAN) 5 October 1989	
P,X	WO,A,9 106 678 (SRI INTERNATIONAL) 16 May 1991 see abstract; claims	1-3, 5-14, 16-22

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. US 9201691 SA 57791

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information. 18/06/92

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0412883	13-02-91	FR-A- 2650840	15-02-91
		AU-A- 6180190	11-03-91
		WO-A- 9102087	21-02-91
WO-A-8909283	05-10-89	US-A- 4971903	20-11-90
		AU-A- 3354889	16-10-89
WO-A-9106678	16-05-91	EP-A- 0450060	09-10-91